The psychology of insurance

van Wolferen, J.

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

THE PSYCHOLOGY OF INSURANCE

Job van Wolferen

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college van promoties aangewezen commissie in de aula van de Universiteit op vrijdag 19 september 2014 om 14.15 uur door Job van Wolferen, geboren op 15 september 1987 te Eindhoven.

Index

Introduction

Insurance is ubiquitous. People insure homes, cars, investments, lives, pets, smartphones, and more. When I started working on this dissertation in September 2010 the total gross premium written in the 34 countries within the Organization for Economic Cooperation and Development (OECD) was $4.306.900.000.000. The latest statistics indicate that number had already grown by 8.5% to $4.673.156.000.000 in 2012 (OECD, 2014). In fact, by many standards people are over-insured: they prefer and buy policies with low deductibles when higher deductibles would save them considerable amounts of money (Sydnor, 2010) and insurance against probable small losses is in high demand (Slovic, Fischhoff, & Lichtenstein, 1977). One particularly excessive insurance policy was available between 2000 and 2006: three sisters paid £100 per year for a policy that would cover the costs of "bringing up Christ if one […] has a virgin birth" (the insurance company would pay £1.000.000 if immaculate conception occured; BBC News, 2006). I think the ubiquity and high global presence of insurance in itself justifies research on topics related to insurance (for estimates and statistics of penetration rates see Swiss Re Ltd, 2013). However, there are two important theoretical reasons to focus specifically on the psychology of insurance decisions.

First, the study of insurance has traditionally been conducted within the realm of economics and finance. In other areas such as the study of pensions (Benartzi & Thaler, 2007) or tax behavior (Kirchler, 2009), knowledge from economics and psychology has successfully fused into behavioral economics to provide new insights into why people decide and act the way they do. Using psychological or behavioral findings is still relatively new in insurance research. A recent book by Kunreuther, Pauly, and McMorrow (2013) synthesizes all the available behavioral economics research on how people and insurance companies make insurance decisions. By comparing observed behavior to benchmark models the book "attempts to understand why […] anomalies sometimes occur and sometimes not" (p.vii). Although my research focuses less on the comparison to formal benchmark models, I follow their lead in using behavioral science research to gain greater understanding of people's insurance decisions.

The second reason for studying the psychology of insurance decisions is that there is uncertainty regarding two topics of (possibly) major importance in how people respond to insurance. First, it is unclear under what conditions (if any) ex ante moral hazard occurs. That is, insurance is predicted to increase the amount of risk people take (for an overview of the theoretical assumptions underlying this prediction see p. 276 in Baker, 1996). In practice however, it is hard to find consistent evidence of ex ante moral hazard across insurance domains. Second, it is clear that insurance fraud is widespread (estimates range from 5% to

20% of premium volume; Coalition Against Insurance Fraud, 2013; Insurance Journal, 2012), but it is not clear *why* people think it is acceptable to commit insurance fraud.

For the two topics described above, it is valuable to obtain greater understanding of when and why people act and feel the way that they do. Such knowledge about the psychology of the insured will add precision to predictions of when insurance induces risk taking, and when not, and under what conditions people think insurance fraud is acceptable. Ultimately, research that integrates findings from both psychology and economics will help design policies that reduce undesirable forms of ex ante moral hazard and aid in the fight against insurance fraud.

In this dissertation I ask three distinct questions related to the psychology of insurance decisions. Specifically, I ask (1) when and why do people opt for insurance?, (2) what is the effect of having or not having insurance on risk perception and risk taking?, and (3) why do people find it acceptable to defraud insurance companies? In my attempt to answer these questions I build on economic, psychological, and other relevant literature and collect data in observational as well as experimental studies. Before continuing to the chapters in this dissertation, it is important to define the concepts around which my research revolves. Therefore, I explain what insurance is, what the major concepts in the insurance literature are, and I briefly discuss theories of individual decision-making that relate to these concepts.

Insurance

Insurance entails incurring a small cost now to avoid potential larger losses later. In most cases, insurance is a tool to reduce financial risk. By paying a small premium people can eliminate the possibility that they have to pay large amounts of money when things go awry. The instrument that makes insurance possible is usually risk sharing; by pooling individual risks a group of people or an insurance company can redistribute individual losses across its members. It is generally agreed upon that the earliest forms of insurance is described in the Code of Hammurabi, which was engraved in a stone around 1772 BC (Harper, 2010). This code of law stated that loans should not have to be repaid if personal catastrophe occurred. The extensive insurance business as we know it today originates in marine insurance and was heavily influenced by the underwriting that took place in Lloyd’s coffeehouse, in the seventeenth and eighteenth century in London (Lloyd’s, 2014). Lloyd’s provided specialized shipping intelligence that allowed underwriters to provide shipping insurance at attractive rates. It also allowed them to deny insurance for endeavors they deemed too risky. From Lloyd’s and other coffeehouses that specialized in marine insurance underwriters started to provide other types of insurance (e.g., home-owner’s insurance that indemnified the cost of rebuilding a house if a fire destroyed it) and insurance eventually became as ubiquitous as it currently is.[1]

Today, insurance is usually provided by companies that pool the risk of many policyholders. The law of large numbers enables insurance companies to fairly accurately observe and predict how much money is needed to accommodate all the claims in their pool, without going bankrupt. Because people tend to be risk- and/or loss-averse the insurance company can charge premiums at a rate higher than what would be actuarially fair. This allows them pay for all the policyholders (legitimate) claims, to incur the costs of running a business, and to make a profit.

Insurance and adverse selection

Since the 1960's and 1970's theories of how people respond to insurance have been proposed, developed, and tested, mainly from within the academic economics literature (for an overview see Loubergé, 2000). One of the major concepts in the literature is adverse selection. It entails that people who are most at risk are most likely to buy insurance: when insurance companies do not or cannot know the amount of risk an individual policyholder runs they are forced to offer insurance at a fixed price. Consequently, insurance is an unattractive deal for the low-risk people but an attractive deal for high-risk people. Therefore, the insurance company is likely to end up with a pool of high-risk policyholders who file many (high) claims. The result is an inefficient insurance market with expensive policies and many uninsured people (Akerlof, 1970; Arrow, 1963; Pauly, 1968; Rothschild & Stiglitz, 1976).

For adverse selection to take place people should have a better knowledge of their risk than their insurance company does. People can only have superior knowledge about their risk if they can accurately assess risk, and there is plenty of literature to suggest that people are generally not so good at doing so. In fact, risk-assessments are subject to biases (e.g., Finucane, Alhakami, Slovic, & Johnson, 2000) and they are often influenced by factors that do not have much informational value. For example, a disaster that can easily be imagined is perceived to be more likely than a disaster that cannot be as easily imagined. As a result, in one study, participants were willing to pay more for $100.000 worth of life insurance that covered deaths due to 'any act of terrorism' than if it covered deaths due to 'any reason' (Johnson, Hershey, Meszaros, & Kunreuther, 1993; for another example see Chapter 3). Although I do not directly measure perceived risk, the findings in Chapter 5 provide a

[1] This paragraph cannot provide a complete picture of the history of insurance. For a more detailed overview see Trenerry (1926) or Bernstein (1998).

demonstration that people do not use the risk-information available to them in a way that adverse selection would predict. Specifically, in a course where students could obtain insurance against failing an exam by scoring a ‘5’, I find that the students who are most at risk are *not* most likely to obtain the insurance.

Insurance and moral hazard

A second major concept in the insurance literature is moral hazard, and there is a distinction between ex ante and ex post moral hazard. Ex ante moral hazard predicts that policyholders take more risk when they are insured compared to when they are not insured because they know the insurance company will cover their losses. The intuition is that people have less reason to be careful if they do not suffer the consequences of their actions. Ex post moral hazard predicts that people are more likely to make use of a service when it is insured compared to when it is uninsured. For example, health insurance reduces the (marginal) price of health care and thereby increases the likelihood that people will visit the doctor when they are sick.

Similar to adverse selection, these two processes could have major consequences for insurance companies; both imply that insurance companies cannot set proper insurance premiums by observing the level of risk-taking in an uninsured population. Such a procedure would lead them to underestimate the actual use of the insured service once their policyholders obtain coverage. Unsurprisingly, research has been focused on identifying the theoretical conditions under which the different types of moral hazard should occur after Arrow (1963) and Pauly (1968) published their paper and comment on the “economics of moral hazard”. However, the progress made in theoretical papers (e.g., Ehrlich & Becker, 1972) has not been matched by empirical tests of its predictions (see Chapter 1 for an overview of the empirical work on moral hazard in health insurance) and this literature could benefit from behavioral science research on whether, when, and why insurance affects risk taking.

In this dissertation I take a couple steps towards a better understanding of the psychology of the insured by studying how insurance affects risk perception (Chapter 2) and risk taking (Chapter 3). I add to the literature that describes how people perceive and respond to risk (Loewenstein, Weber, Hsee, & Welch, 2001; Risen & Gilovich, 2008; Slovic & Peters, 2006; Slovic, 1987). Specifically, my dissertation was given impetus by the finding that reminders of insurance lead to lower risk assessments (later named "the insurance effect"[2]; Tykocinski, 2008; 2013). In a study that was conducted among train commuters in Israel, half the participants were reminded that they had health insurance. The other half was not. The reminded people consequently thought they were less likely to need medical care in the next five years than the non-reminded people. My co-authors and I hypothesized that decreases in the perceived likelihood of misfortune could explain why people with insurance take more risk than people without insurance. In Chapter 2 I therefore replicated the study in two Dutch samples and one American sample. I found no support for the 'insurance effect', but did find evidence for tempting fate effects: I replicated the finding that people think they are more likely to be called upon in front of class when they know they are not prepared (Risen & Gilovich, 2008). Together, the findings in Chapter 2 suggest that people are more strongly influenced by being uninsured than by being insured.

In Chapter 3 I build on these findings and conduct controlled tests of ex ante moral hazard predictions. As can be read in the introduction to that chapter, but also in Chapter 1, there is quite some uncertainty about the existence and the possible size of ex ante moral hazard effects. Consistent with the results in Chapter 2, I find that *not* having insurance has the largest effects on people's (hypothetical and actual) willingness to take risk.

Insurance fraud

Insurance fraud is the final 'concept' that could benefit from research that integrates knowledge from multiple social and behavioral sciences. Insurance fraud means that people file claims for damages that did not occur or that they misrepresent the nature of non-covered damages so that the insurance company will cover them. A simplified model of insurance fraud states that people decide whether or not to commit insurance fraud based on a cost-benefit analysis of all the possible positive and negative consequences of committing fraud (Becker, 1968). Research on when, why, and how people cheat proposes that people cheat to the extent that they can justify the lie to themselves (Mazar, Amir, & Ariely, 2008). There are factors that facilitate the justification of lies and it turns out several features of insurance provide strong opportunities for people to justify insurance fraud. First, insurance companies are large and anonymous victims. Research on the identifiable victim effect (e.g., Jenni & Loewenstein, 1997) suggests that this is the exact opposite of what would be needed to prevent insurance fraud. Second, based on simple expected-value calculations, insurance is

[2] The term insurance effect was coined in Tykocinski (2013) and I use it in the introduction and discussion of my dissertation. Before 2013, I described the finding that insurance makes misfortune seem less likely as the "protection effect" (see Chapter 2). However, both terms mean the same thing.

almost never a good deal. Equity theory (Walster, Hatfield, Walster, & Berscheid, 1978) and the norm of reciprocity (Gouldner, 1960) dictate that if one party harms the other, people are motivated to restore the balance between the two. If people feel the insurance company is taking advantage of them, they may feel justified to take advantage of the insurance company in return. In this dissertation I describe an additional feature of insurance that allows people to justify insurance fraud to themselves. Namely, in Chapter 4 I find that most people do not readily think of the sharing aspect of insurance. Instead, people are inclined to think about insurance as an individual contract. Fraud is then seen as taking from a rich and anonymous company, and not as taking from the fellow insured. People think that they should get some money back from their insurance company in return for the premiums that they paid. In that same chapter, I propose that these ideas and feelings facilitate accepting attitudes towards insurance fraud.

Note

Before continuing one should know that all chapters in this dissertation could be read individually and in any order. That is because they constitute separate articles that are or will be published individually. Hence, there might be some overlap between the introductions of the different chapters. In addition, I wrote the first and final chapter using 'I' but in all the other chapters I use 'we'—the reason is that the other chapters were coauthored by my supervisors.

Chapter 1 - Moral hazard in the insurance industry

Abstract. This Panel Paper reviews recent evidence on moral hazard in the insurance industry. We discuss three types of moral hazard and detail how each is an asymmetric information problem. For each of the types, we summarize the empirical evidence and discuss the policy implications that follow from it. The evidence for ex ante moral hazard (i.e., insurance-induced increases in risk-taking) is rather weak but suggests that people engage less in preventive behaviors that are costly and hard to maintain when they obtain insurance. The evidence for ex post moral hazard (i.e., insurance-induced increases in usage of insured services) overwhelmingly indicates that it exists. However, the exact size of this effect in the Dutch health care system remains to be empirically determined. The numbers on insurance fraud indicate that it poses a significant problem. Furthermore, insurance fraud is deemed acceptable and common by many policyholders, and most seem unaware of the nature of insurance. Despite the large amount of data that has been gathered and analyzed, policymakers often lack the knowledge required for accurately predicting what effect policy changes will have. An important implication that follows from this is that insurance companies need to collect more data to gain the knowledge they need.

This chapter is published as: Van Wolferen, J., Inbar, Y., & Zeelenberg, M. (2013) Moral hazard in the insurance industry. *Netspar Panel Paper, #33,* 1-73.

In the past five years, health care spending in the Netherlands has continued its upward trend. Going from €70.7 billion (13.1% of GDP) in 2006 to €90 billion (14.9% of GDP) in 2011, the demand for health care places an increasing burden on the Dutch economy and its citizens (CBS, 2012). A considerable share of the increased demand is likely due to the aging of the Dutch population, and could have been foreseen. The combination of this trend, however, with the current financial climate, has led to an unprecedentedly loud call for budget cuts in the health care industry.

The budget cuts are also part of the Dutch political discussion, as health insurance coverage is mandatory under the Dutch law. Politicians thus decide which treatments should (at least) be covered under the universal policy, and which should not. Although it is clear that political parties disagree *which* budgets should be cut, they all agree that spending should be reduced. One important factor in deciding which budgets to cut is to determine where money is spent unnecessarily. While much of the discussion centers on content—how valuable given treatments for certain diseases are, how much of care for the elderly should be covered—we focus on a more universal problem that leads to overspending. We discuss three different types of moral hazard and explain how each of them contributes to the increasing burden on the Dutch health care system.

For each of the three types, we start with an explanation of how that type of moral hazard affects the behavior and/or demand of a given insured person. Then, we review the available empirical evidence and discuss what those findings imply for policy and policymakers.

What is moral hazard?

Moral hazard broadly refers to the increase in people's use of a service when it is covered by insurance, compared to when it is not. We will use doctor-patient situations and related health risk examples at different points in this paper to illustrate what specific theories or predictions entail. A general example would be that when people have health care insurance they are more likely to visit a doctor compared to when they do not have health care insurance.

Moral hazard theory describes what happens when one party is partly insulated from risk because another party agrees to wholly or partly indemnify losses that the first party might suffer. In a health insurance context, for example, everyone is at risk of needing medical care at some point. This care, of course, costs money, and the insurer agrees to pay for it provided that the insured pays a monthly premium. The main point of moral hazard theory is that insurance removes all or part of the incentive to restrict the use of insured services. With full insurance (i.e., 100% of the costs are covered) there is no financial reason for someone with insurance to not visit the doctor.

The presence of asymmetric information underlies the moral hazard problem (Arrow, 1963). The asymmetry occurs because the insurer has less information about the health status and reasons for health care usage of the insured than the insured themselves do. The insurer cannot check whether a physician visit would not have been made without insurance (when the visit would cost the insured more), or because of some avoidable risks that the insured took. The insured may even have faked his doctor visit to obtain money from the insurance company. In each of these examples, the insurer is unable to perfectly observe the risks that the insured takes and adjust the contract accordingly. It is important to note that these risks need to be under the control of the insured. If they are, moral hazard may occur: The insurer prefers that the insured take precautions to avoid risks, but cannot enforce this wish. The insured is able to take risks without this being priced into the insurance contract and this asymmetry may lead to greater usage of insured services.

Large theoretical frameworks elaborate on the market failures that may occur as a result of asymmetric information between insurer and insured (Akerlof, 1970; Arrow, 1963; Ehrlich & Becker, 1972; Pauly, 1968; Rothschild & Stiglitz, 1976). These market failures include the complete eradication of insurance but also more subtle forms where asymmetric information leads to sub-optimal levels of insurance. Empirical tests of moral hazard theory in the insurance industry have become more prevalent recently, and three types of moral hazard dominate the discussion.

Different types of moral hazard

The different health care examples described above hint at the fact that moral hazard may occur for different reasons. This section describes ex ante moral hazard, ex post moral hazard, and insurance fraud. These are theoretically, psychologically, and behaviorally distinct ways in which health insurance might increase the use of medical services. It should be noted that these types of moral hazard also apply to different domains.[3] We focus on health insurance specifically, as this is a domain in which moral hazard theory has spurred much discussion (Savedoff, 2004) and where it may have far-reaching practical consequences.

To illustrate why the first type of moral hazard is predicted, we present a formula on the basis of which people are predicted to determine their actions:

[3] For example, workers compensation (Dionne & Michel, 1991), auto insurance (Chiappori & Salanie, 2002; Richaudeau, 1999) and even pensions (Keating, 1991).

$$Expected\ loss = p \times L$$

where p represents the probability or likelihood of the loss whereas L represents the size of that loss. Running the risk of over-explaining: if the risk of getting sick during a given period is 5% (p) and it costs 100 Euros to visit the general practitioner (L), the expected loss is .05 × 100 = 5 Euros. Economic theory assumes that people are rational utility maximizers. This means that they will try to find the optimal solution to this formula. Let us stick with an expected loss of 5 Euros. For the purposes of this example, we will not specify a full model but will simply assume that the expected loss of 5 Euros is an optimal solution for a given insured person named Jeff. Assuming Jeff is a rational utility maximizer, he should change his behavior when p or L changes to return to the—optimal—expected loss of 5 Euros.

Now look at what happens when health insurance is introduced. By compensating for the costs of visits to the general practitioner (GP), insurance basically reduces L. If insurance covers 90% of the costs of the visit, L becomes .10 × 100 = 10 Euros—the amount that Jeff will need to pay the doctor. If insurance by itself does not change the likelihood that Jeff gets sick (p) the expected loss becomes .05 × 10 = 50 Eurocents. Jeff—the rational utility maximizer—does not like this suboptimal expected utility and will change his behavior to return to the optimal level of expected utility.

Behaviors that are referred to as *ex ante moral hazard* change p and take place only before (hence, *ex ante*) an incident takes place. By taking more risk, people increase their expected loss. For example, Jeff used to take vitamin pills and therefore only had a 5% chance of getting sick (the incident). Now that he is insured, he stops taking his vitamin pills and thereby increases his chances of being in a position to need a doctor. Without vitamin pills, the likelihood that Jeff needs to see a doctor is 50% (they were effective pills, indeed): .50 × 10 = 5 Euros. So by refraining from using preventive care, Jeff returns to his optimal level of an expected loss of 5 Euros and thereby maximizes his utility. This negative effect that insurance has on preventive health behaviors is predicted to hold across behaviors that positively or negatively affect p, including smoking, exercising and alcohol consumption.[4]

The second type of moral hazard—*ex post moral hazard*—relates mainly to behaviors that change L and take place only after (hence, *ex post*) an incident has taken place. By using

[4] A notable theoretical exception is the prediction by Ehrlich and Becker (1972). They mathematically examine the case where an insurance company observes risk-taking and prevention efforts perfectly, and prices insurance contracts accordingly. In this case, policyholders face an incentive to be careful when insured, as carefulness lowers their premium. This is such an unlikely assumption that we do not discuss its practical implications further.

more of an insured service, people again try to maximize the utility they experience. Considering Jeff again, imagine that he is sick. Remember that without insurance Jeff pays 100 Euros for a GP visit. Given that Jeff has a limited amount of money to spend, he has to decide whether visiting the doctor is worth 100 Euros to him. If Jeff were seriously ill, he would probably think the visit is easily worth the price. If he just had a cold, however, he might decide to wait and see if the cold passes without visiting a doctor. In both cases, the rationality assumption implies that Jeff determines whether the utility he (thinks he) gets from a GP visit outweighs the costs.

Now let's look again at what happens when Jeff has insurance. Suddenly, visiting the doctor costs 10 Euros. To Jeff—who thinks that 100 Euros is too much to cough up for a cold-related GP visit—10 Euros might seem like a fair price. With insurance, Jeff thus visits the doctor when he has a cold. Without insurance, he does not. By reducing the price of care, insurance is thus hypothesized to increase the demand for care.

In a related yet different scenario, imagine that when Jeff has a cold, he is willing to spend 100 Euros on a visit to the physician. When insurance reduces the price of the visit to 10 Euros Jeff has 90 Euros 'left' to spend on care. In order to maximize his utility, Jeff might demand additional tests from his doctor or make additional appointments that he does not necessarily need.

To someone who is not trained in economics it might seem strange to describe people's choices as a function of some utility calculations. Indeed, why would someone who is insured stop taking vitamin pills or see the doctor more often when sick? Modeling human behavior as a calculation-based enterprise is central to economic theory. Throughout the paper, we discuss the benefits and limits to economic theory and how its specific rationality assumptions sometimes prove hard to maintain when describing the behavior of irrational people.

The final phenomenon that resides under the moral hazard umbrella is insurance fraud. Unlike the other two types of moral hazard, fraud does not increase the demand for care. However, given that insurance fraud wouldn't be possible without asymmetric information—and cheating on insurance companies is deemed immoral—it is referred to as a moral hazard. The asymmetry of information comes into play because the insurer is unable to directly observe why the policyholder files a claim—so whether the claim reflects genuine or faked use of insured services.

The simplest theory of insurance fraud states that a policyholder commits fraud when the benefits of committing fraud outweigh the expected loss. The expected loss is determined as the (subjective) probability of being caught, times the punishment that follows being caught (e.g., fines, exclusions, jail time). Often, rather than modeling the decision to commit insurance fraud, the propensity to commit fraud among policyholders is *assumed* to be non-negative. Much of the scientific literature on insurance fraud focuses on how fraud can be detected or what types of incentives should be built into insurance contracts to prevent insurance fraud.

Returning to the calculative Jeff: when he thinks he can get away with claiming reimbursements for a treatment he never received, he will do so if that is profitable enough. Similarly, when signing up for insurance, Jeff presumably knows his risks better than the insurer can observe them (e.g., family history of diseases, smoking and alcohol consumption, how often he exercises etc.). Jeff might hide some information in order to get a lower premium or greater coverage—which would have been denied him, had he shared his knowledge with the insurer. In both cases, Jeff uses the informational advantage he has over the insurer to his own benefit.

A short note on the morality of moral hazard

Before we consider the empirical evidence that has been gathered to test moral hazard theory, we briefly discuss the morality of moral hazard. Historically, economists and the insurance industry have disagreed on whether increased usage of services as a result of moral hazard is immoral.[5] Economists regard insurance-induced changes in behavior as rational and do not have a strong opinion about these changes—if anything, they consider rational behavior to be good because it leads to 'optimal' outcomes. In contrast, insurers, unsurprisingly, have considered increased use of insured services as immoral. Their reasoning is that people should not, simply because it's free, use something that they apparently do not consider to be worth its cost. We consider insurance fraud to be wrong and immoral because it tarnishes the fundaments of insurance. It is hard to condemn ex ante and ex post moral hazard similarly strongly. We will see that some instances of moral hazard are desirable, while others are clearly not. We thus refrain from taking a clear stand in this discussion and continue with describing the extent and strength of the evidence for each type of moral hazard.

[5] We refer interested readers to (Baker, 1996) and (Rowell & Connelly, 2012) for in-depth analyses of how moral hazard got its name and how theorizing has changed over time.

Empirical tests of moral hazard theory

This section discusses the available studies according to the type of moral hazard they test. We start with ex ante moral hazard, continue with ex post moral hazard, and conclude by discussing insurance fraud. Each section opens with an explanation of what the specific type of moral hazard might look like in a health care context, followed by a discussion of the available empirical evidence, and closes with a discussion of policy implications. However, before a start can be made, we have to explain the biggest problem in empirical tests of moral hazard in real-world data: disentangling moral hazard from adverse selection.

Adverse selection

Asymmetric information is conducive to both moral hazard and adverse selection. Adverse selection means that people with high risk are more likely to opt for insurance than people with low risk. In health care, people who have a family history of sickness and disease are more likely to buy health insurance than people whose family members have lived long and healthy lives.[6] Insurers cannot perfectly observe a person's health and are thus unable to adjust the insurance premium according to the expected health care expenses.

Although adverse selection has to do with the demand for health insurance, while moral hazard concerns behavior caused by insurance, both lead to the same outcome. When unhealthy people are more likely to buy insurance, insured people will turn out to be unhealthier than uninsured people. As a result, the demand for health care will be greater for insured people than for uninsured people.

If insurance induces moral hazard and causes the insured to excessively use health care, the insured also demand more health care than the uninsured. If we wish to understand how insurance affects people's health decisions it is thus not sufficient to merely test whether the insured demand more health care than the uninsured. Such a comparison would leave it unclear whether the insured demand more health care because they were sicker than the uninsured in the first place, or because insurance lowered the cost of health care and thereby increased demand.

The scientific community has developed several ways to estimate the size of each effect independently (e.g., instrumental variables, natural experiments, longitudinal studies, or response to changing incentives within insurance plans; for an intruiging example, see

---

[6] Note that adverse selection is different from insurance fraud. The former is merely the result of a correlation between health status and preference for insurance. The latter is unrelated to the health status of the prospective policyholder and implies that he is consciously shielding information from the insurer to obtain personal gain.

Abbring, Chiappori & Pinquet, 2003). However, none of these are as good as randomized control trials— and we will see that different estimation strategies sometimes lead to different estimates. In the following sections it will become clear that uncertainty remains with respect to the size of the ex post moral hazard effect and the prevalence of insurance fraud. This highlights the need for randomized control trials in the insurance industry.

Ex ante moral hazard

This section takes a look at the empirical tests of ex ante moral hazard to find out whether insurance leads people to take more risk and exert less prevention effort. In the context of health care insurance, the studies mentioned here thus test whether Jeff—upon signing an insurance contract— would indeed stop taking his vitamin pills, start drinking and smoking, or stop exercising.

Empirical evidence for ex ante moral hazard

Most work on ex ante moral hazard has been theoretical and there have been only a few empirical investigations into the matter (Kenkel, 2000; Zweifel & Manning, 2000). We review the articles that report empirical investigations of the ex ante moral hazard effect of health insurance.

The first study that we discuss here (Stanciole, 2007) analyzes data on health behaviors between 1999 and 2003 from the U.S. Panel Study of Income Dynamics (PSID). The PSID surveys close to 8,000 families, collecting data on a range of topics—including whether or not people have health care insurance, how healthy they are, and to what extent they engage in a healthy lifestyle. Stanciole (2007) specifically focuses on heavy smoking, heavy drinking, sedentarism (i.e., never engaging in light physical activity), and obesity.

A simple comparison of the lifestyle of insured (approximately 93% of the sample) and uninsured people would lead one to conclude that the uninsured live unhealthier lives than the insured—a case of reversed moral hazard! However, we have seen that a simple comparison between the two groups is not sufficient. The richness of the data provided by PSID allows the researcher to try and filter out many factors that affect people's lifestyle choices and insurance decisions (e.g., income, education, health, area of residence, and so forth). Such procedures make it more likely that the true effect of insurance on lifestyle choices is extracted. Using an elaborate econometric model, Stanciole (2007) finds that being insured is associated with a greater likelihood of heavy smoking, sedentarism, and obesity, but less heavy drinking—controlling for the many factors mentioned above. The author then concludes that insurance causes policyholders to engage in unhealthy behaviors. However, the finding is a simultaneity rather than causality. Even though 93% of the sample was insured, and despite the attempt to filter out any selection effects, the presence of adverse selection cannot be ruled out.

A similar study uses data from 6,000 households (approximately 15,000 individuals) stored in the British Household Panel Survey on smoking behavior and the answer to a yes/no question ‘do you play sports, go walking or swimming at least once a year’ (Courbage & de Coulon, 2004). Approximately 10% of the population opts for private insurance coverage in addition to the publicly funded national health care system. The effect of this additional coverage on ‘preventive’ health behaviors is estimated. Note, however, that the measures were taken from a dataset that was not specifically designed to test for ex ante moral hazard, which means that the measures of preventive behaviors are therefore rather crude.

The data do not support the ex ante moral hazard prediction. In fact, people with private insurance coverage are more likely to exercise and are less likely to smoke. The authors carry out an additional test, estimating the effect of private insurance on the likelihood that women undergo a breast cancer screening or seek cervical smear testing. Both are preventive behaviors and the ex ante moral hazard hypothesis thus predicts that insurance decreases the likelihood of these screenings. The complication, however, is that these preventive behaviors are covered by the insurance policy. The ex post moral hazard would thus predict an increased use of these services among the privately insured—and indeed a small positive effect for breast cancer screening is observed (Courbage & de Coulon, 2004).

A limitation to this study is that it controls for far fewer factors than the previous one (Stanciole, 2007) did, and it suffers from a similar selection problem. That is, the authors cannot exclude the possibility that people opted for insurance because they were unhealthier than the people who did not opt for insurance. With observational datasets such as these, causal effects may be estimated using ‘instrumental variables (IV)’. We will not go into the details of this technique, but it is useful when randomized control trials are impractical or impossible. However, the IV estimate also shows that private insurance does not lead to less exercising or more smoking. In sum, this study thus contradicts the ex ante moral hazard prediction.

A very specific and interesting test of the ex ante moral hazard effect was conducted after different U.S. states—at different times—adopted mandates that required basic insurance policies to cover “supplies, services, medications, and equipment for treating diabetes […] without charging higher premiums” (Klick & Stratmann, 2007). If insurance coverage of diabetes treatments causes diabetics to reduce their prevention efforts (i.e., watching their diet) then their Body Mass Index (BMI) should increase. The researchers use BMI data from the Behavioral Risk Factor Surveillance System accumulated between 1996 and 2000, and tested whether BMI went up for diabetic individuals after the state in which they lived adopted an insurance mandate. Astonishingly, they find that when states adopt such mandates the average BMI of diabetics increased by 1.7 points, or 6 percent! Importantly, the BMI of non-diabetics did not respond to the mandates. It thus seems that diabetics trade prevention for treatment, possibly because prevention is so costly. Namely, prevention of diabetes worsening is strenuous work—it is hard to avoid sugary foods and to stick to a healthy diet. Furthermore, the treatment of diabetes-related problems is effective. It is thus unclear to what extent this finding generalizes to other specific diseases and treatments for which prevention efforts are less costly or treatment is less effective.

Recently, another highly specific test of ex ante moral hazard was conducted on data obtained in a malaria-prone area in Ghana (Debebe, van Kempen & de Hoop, 2012). There, insurance for malaria treatments was introduced and the authors tested whether this led to a reduction in six different malaria prevention activities (e.g., repellent use and expenditure, presence of mosquito-proof windows, use of mosquito nets). Interestingly, insurance for malaria treatments only reduced the use of mosquito nets that have to be maintained by the owner—a demanding and time-consuming task that involves treating the net with insecticide. For 'factory-nets'—that do not require treatment with insecticide by the owner—no reduction in use was observed. This finding, together with the Klick and Stratmann (2007) study, indicates that people reduce their preventive efforts only to the extent that it is hard to engage in them.

By far the best and least confounded test of whether insurance leads to less healthy behavior was conducted by Dave and Kaestner (2009). In the U.S., Medicare provides the opportunity to everyone over who is 65 years of age or older to enroll in an insurance plan. Compared to private insurance plans, Medicare is very generous and cheap. In fact, part A (hospital insurance) is free and the optional part B (medical insurance) costs very little per month. Part C involves the option to receive Medicare benefits through their private insurance plans and part D deals with prescription drug plans (both are optional). Unsurprisingly, the participation rate in plan A and B among people eligible for Medicare was 97.13% in 2008 (New Geography, 2008; State Health Facts, 2012). Dave and Kaestner (2009) describe the provision of Medicare as causing an 'exogenous' shock in health care insurance. Since virtually all adults over 65 opt for Medicare there is no selection problem in testing the effect of insurance on health behavior. That is, people do not choose to (not) enroll in insurance; they are 'assigned to insurance'. The researchers exploit this situation to estimate the pure effect of insurance on health behavior using data of 3,396 people from the Health and Retirement Study, collected in eight measurement waves. These data are specifically collected for testing hypotheses such as moral hazard and thus include a vast amount of information at the individual level.

The authors note that physicians generally encourage healthy lifestyle choices and discourage unhealthy alternatives. Confirming the ex post moral hazard hypothesis, obtaining Medicare increases the number of visits to the physician—thereby leading Medicare enrollees to live healthier lives. The authors then cleverly control for the positive effect of increased physician contact on health behavior to estimate the true effect of insurance on preventive behaviors.

Even though this study avoids the selection problem, it has to deal with the fact that many people have insurance before they are eligible for Medicare. This means that they can only test the effect of insurance on people who have never had insurance prior to obtaining Medicare. This narrows the population under investigation drastically and introduces a confounding variable. People who have never been able or willing to buy insurance are likely to differ from people who have had insurance in many ways. The authors attempt to control for these differences by filtering out the effects of age, sex, race, employment status, wealth, and so forth, but it remains hard to determine how well they capture pre-existing differences between the enrollees that have and those that have not had insurance.

Controlling for the positive effect of increased contact with physicians, an ex ante moral hazard effect is found for men (who had never been insured prior to obtaining Medicare). Most notably, when those men enroll in Medicare, they become almost 40% less likely to exercise, 18% less likely to stop smoking, 15.8% more likely to smoke daily, and 14.8% more likely to drink daily. No such effects were observed for women who enrolled in Medicare. More importantly, the positive effect of increased contact with physicians on these behaviors fully counteracted the negative effects of insurance. The net effect of insurance on health was therefore predicted to be zero. In conclusion, ex ante moral hazard is of little concern in the Medicare system—the ex ante moral hazard effect that is found occurs for only a small subset of the insured, and it is not clear that the finding generalizes to the entire Medicare population.

Although the Dutch health insurance system is rather different from some of the systems in which the studies above have been conducted, we now elaborate on how the described findings could inform Dutch policymaking.

Policy implications of the findings on ex ante moral hazard

Unfortunately, the empirical evidence thus far does not paint a clear picture of whether, when, and why ex ante moral hazard occurs. One obvious ‘policy’ implication is that insurers should work together with academics to come up with new ideas to test the *ex ante moral hazard* hypothesis, while using data from insurance companies. This will advance both theoretical and practical insights into whether, when, and how people’s risky decisions are affected by insurance. We highlight the importance of small-scale experiments again later in this section.

Despite considerable variability in estimates of the size of the ex ante moral hazard effect at least two conclusions can be drawn. First, insurance does not seem to cause excessive risk-taking. Rather, ex ante moral hazard effects are only found for preventive behaviors. The findings would thus allow us to predict that Jeff is less likely to take his vitamin pills when he obtains insurance— but not that he would start to smoke or use drugs. Namely, even the effects in studies that test for perverse effects of insurance on drinking behavior and exercising might reasonably be construed as reductions in preventive behaviors: Upon enrolling in Medicare, elderly men were less inclined to stop or reduce their smoking and alcohol intake.

Second, the empirical evidence suggests that ex ante moral hazard is most likely to occur for diseases where people have to take many medications, follow a strict diet, or exercise a lot (Debebe et al., 2012; Klick & Stratmann, 2007). These behaviors are very ‘costly’, in that they require a lot of time, effort, and devotion on the part of the insured. To give some concrete examples, asthma patients are required to use their asthma pumps twice a day, every day; people with heart problems have to exercise regularly to stay fit; and people with high blood sugar have to follow restrictive diets. Such preventive efforts are hard to maintain, and the reviewed evidence suggests that these behaviors are especially susceptible to ex ante moral hazard effects. Insurers that cover the treatments for these diseases would therefore be well-advised to look into procedures that enforce people’s adherence to these preventive measures. The reasonable thing to do for companies that are trying to promote these healthy behaviors is thus to eliminate or at least try to reduce some of the hassle associated with preventing the disease. In the case of asthma, companies should allow their clients to sign up for a daily e-mail or texting service that reminds them to take their medication. Similarly, people with heart problems could be encouraged to define an exercising schedule or to develop a routine. Adherence to the schedule could be enforced in a similarly simple fashion. These are relatively cheap and likely effective measures to mitigate ex ante moral hazard.

The studies described in this section of the paper also provide guidelines that help in determining how much policymakers should worry about ex ante moral hazard when new treatments become available. In deciding whether or not to cover the treatment, and to what extent incentives to counter ex ante moral hazard should be provided, it would help to determine how much trouble it is for people to engage in preventive efforts. If taking precautions against a specific disease is easy, it is unlikely that providing coverage for the treatment of that disease diminishes the extent to which people engage in them. In contrast, if people have to exert a lot of effort to prevent a given disease, they are more likely to be responsive to the incentive effects of insurance.

Looking at the current literature, the findings suggest that even if there is an ex ante moral hazard effect, it is probably rather small. In addition, it is likely that the net effect of insurance on preventive efforts is zero. This follows from Dave and Kaestner (2009) who find that insurance also leads to healthier lifestyle choices through increased contact with physicians. It is important to note that this does not mean that there is no reason to try to prevent ex ante moral hazard. Namely, the net effect of insurance might turn out to be positive if insurers could prevent the decrease in healthy lifestyle choices and/or preventive efforts while maintaining the positive effect of increased contact with physicians.

We end this policy recommendation section with a word of caution and a call for experiments. We were unable to find tests of ex ante moral hazard in the Dutch health care system and it is unclear whether the findings in the studies we described are directly applicable to the Dutch system. In contrast to some of the systems that have been studied, health care insurance is legally required and thus nearly universal in the Netherlands. Even if there were a test of ex ante moral hazard in the Dutch system, there would be no uninsured group to compare the Dutch behavior to. This would cause difficulty with estimating the ex ante moral hazard effect.

One possibility is to look at how different *levels* of coverage affect people's health decisions. Then, however, the problem of disentangling adverse selection from moral hazard arises. This highlights the need for insurers to engage in small-scale experiments in which they randomly vary the coverage for certain treatments for some (randomly selected) group of policyholders. Randomized control trials are the golden standard in scientific research and are the best method to obtain reliable estimates of the effect under study. In the case of ex ante moral hazard, an example would be to increase diabetes treatment coverage for a randomly selected group of diabetics and non-diabetics, and then measure their BMI and how well they stick to their diet. Our guess is that few policyholders would decline greater coverage at the same price and there are nearly limitless variations on this experimental design that could be applied to test predictions in specific domains of interest. Running experiments is the best way to gather the knowledge that is needed to inform policy.

Ex post moral hazard

This section reviews the empirical tests of *ex post* moral hazard. Holding preventative behavior constant, does insurance lead to an increased demand for care? Again, we specifically focus on health insurance, and these papers thus test whether demand for care increases when insurance lowers the price of care. In contrast to ex ante moral hazard—which is mainly perceived as malignant—ex post moral hazard may be construed as both malignant and benign. The malignancy argument holds that insurance induces people to use care that they otherwise would not have bought for the same price. It thus leads to a welfare loss because the consumption of care costs more than its real worth to the consumer. The benign argument holds that being uninsured leads people to forego even basic and necessary care, which makes some increased use of care under insurance appropriate (Nyman, 2004). In addition, especially if care is cost-effective and the treatment does not have negative side effects, more moral hazard is better (Pauly & Held, 1990).

Most ex post moral hazard tests use existing datasets and exploit some sort of exogenous variation in the price of care—for example, when co-payment rates for physician services rise in one group of people but not in the other (Chiappori, Durand & Geoffard, 1998), when turning 65 makes U.S. citizens eligible for Medicare (Dave & Kaestner, 2009), or when co-insurance rates change nationwide (Cockx & Brasseur, 2003). It is then determined how much of the change in demand for care is attributable to these changes in price. Unfortunately, many of these studies are unable to directly estimate the causal effect of the price change on the change in demand for care. This might be due to the absence of an appropriate control group, the cross-sectional design of the study, or the inability to observe the 'unobservables' that interact with the price change and the demand for care. Consequently, a review that includes multiple studies that approach the ex post moral hazard test differently shows that the estimates of the ex post moral hazard effect are somewhat inconsistent. However, most studies conclude that demand for care goes down when the price of care goes up—but only by a bit (Skriabikova, Pavlova & Groot, 2010).

Another problem that troubles research on ex post moral hazard is the limited availability of data. Selecting and finding the right information is frequently a difficult undertaking. Most insurance companies do not keep track of the health behaviors of their beneficiaries and they often have limited information on their backgrounds. In addition, data may be inaccessible for researchers because of privacy laws. Many articles use data from nationwide databases or very ‘local’ databases such as data from one particular insurer. These have the problem that they may not contain all the information the researcher needs to answer his questions (e.g., only ‘proxies’ for demand and health behavior) or the generalizability of the sample under study may be questioned.

Empirical evidence for ex post moral hazard

On July 1st, 1993, a law was passed in France that increased the co-payment percentage in the public health care plan. Most for-profit insurers that provided supplementary coverage responded to this law by also introducing cost-sharing in their plans. Importantly, some of these plans were paid for by firms on behalf of their employees (ruling out adverse selection effects) and some firms decided not to implement the cost-sharing. This situation provided an interesting natural experiment where the employees of one firm kept their insurance plan without cost-sharing, while employees of a similar firm suddenly faced 10.4% cost-sharing for physician services (Chiappori et al., 1998).

In testing the ex post moral hazard hypothesis the researchers distinguish between office visits, home visits (i.e., the physician goes to the home of the sick), and specialist visits. A comparison was made between the amount of visits per person in the year before the implementation of the law and in the year after it. The number of office and specialist visits remained relatively constant in both groups. However, the average number of home visits dropped by 18.75% only among employees of the firm that introduced cost-sharing.

Chiappori et al. (1998) conclude that small monetary costs may not induce large changes in the demand for care because of significant non-monetary costs associated with office and specialist visits (waiting time, etc.). Specifically, they state that the relative size of the monetary cost is not large enough to cause changes in these visits. An alternative explanation could be that ‘luxury’ care (i.e., having the doctor visit you, rather than the other way around) is more sensitive to price changes than basic care. In addition, the beneficiaries were not randomly assigned to their insurance plans. Therefore, it may be questioned whether the employees of firms that introduced cost-sharing are directly comparable to those of the firm that did not.

A study conducted in the U.S. also distinguishes between different types of care and also finds differences in the extent to which ex post moral hazard applies to them. The data come from the Medical Expenditure Panel Survey and all doctor visits in the year 2002 are coded to be either a general physician visit or a specialist visit. In addition, for every visit it was determined whether it was related to an acute, chronic, or no disease (Koç, 2010). The specialist visits may include episodes related to anything ranging from physiotherapy (no disease) to respiratory or heart problems (acute or chronic). Elaborate statistical techniques are used to address the endogeneity problem of insurance—people are not randomly assigned but rather choose to be publicly, privately, or not insured at all. Therefore, it remains troublesome to estimate the true causal effect of insurance on the demand for care.

Analyses on the 17,419 observations (of people between the ages of 18 and 64) show the highest moral hazard estimate for specialist visits that are unrelated to diseases. Demand for these visits for those with private insurance is 231% that of the uninsured. In comparison, this number is 'only' 139% for specialist visits related to chronic diseases. There is much less variability in the moral hazard estimate for general practitioner visits across the different types of visits (104% to 151%). This finding makes intuitive sense, as the visits with highest moral hazard are the ones most easily deferred (or perhaps even unnecessary), while the visits with lowest moral hazard are hardly deferrable and most likely necessary. If cost-sharing reduces moral hazard, a policy implication that follows from this (and the Chiappori et al. (1998) finding) is that the level of cost-sharing should be different for different visits. However, while Chiappori et al. (1998) did not find an ex post moral hazard effect for specialist visits, Koç (2010) did.

In the ex ante moral hazard section we described the study that exploited the exogenous variation in health care due to Medicare eligibility (Dave & Kaestner, 2009). Both males and females who were uninsured prior to Medicare had more doctor visits, as well as an increased likelihood of visiting a doctor once and having a hospital stay upon receiving Medicare. That finding is in line with the ex post moral hazard hypothesis.

Recall that ex ante moral hazard predicts that people engage less in prevention when they have insurance. Ex post moral hazard predicts that people demand more care. So what happens when the prevention of a given disease is a form of care? Consider preventive screenings such as breast or prostate examinations, vaccinations, and other prophylactic measures. These treatments might reasonably be considered to be preventive efforts and thus, following from research on ex ante moral hazard, one would expect lower demand for these treatments (i.e., less prevention) among people who have insurance compared to those who do not. In contrast, it is equally defensible to consider these treatments as instances of care and therefore predict that, consistent with ex post moral hazard, the demand for these treatments would be higher for people who have insurance compared to those who do not.

The empirical evidence is in favor of the latter hypothesis—that is, greater use of preventive care by the insured. This evidence comes from a study conducted using data from

12,100 Mexicans over the age of 50 in the Mexican Health and Aging Study (MHAS) to test for differences between the insured and uninsured in the extent to which they use preventive services (Pagan, Puig & Soldo, 2007). This study (unlike Dave & Kaestner, 2009 and Stanciole, 2008) does not find differences in smoking and drinking behaviors. However, a positive effect of insurance on the use of insured services is found. The uninsured in the sample are less likely to use preventive screenings for hypertension, high cholesterol, diabetes, and three types of cancer. This finding is not very strong, as only the effect of insurance on cholesterol and diabetes screening remains significant after controlling for other health risk factors. Nevertheless, this study definitely provides no evidence for reduced use of preventive medical care, whereas there is some evidence for increased use.

It is clear from the studies described above that the ex post moral hazard hypothesis has truth to it. It is fair to say that the demand for care is responsive to the price of care. When insurance reduces the price of care, demand goes up. However, many studies point out that it is hard to draw firm conclusions from the data they present because their nature remains correlational. Next, we discuss a study that circumvents this problem by randomly assigning families to different insurance plans. However, before we do so, we explain a common metric to measure the size of ex post moral hazard: elasticity of demand. Readers with a background in economics may feel free to skip this section.

Demand elasticity

Although not every paper on ex post moral hazard estimates the price elasticity of the demand for health care, it is a very important number in policymaking. Demand elasticity gives insight in the extent to which demand for health care changes when the price goes up. The price elasticity of demand for a certain good is calculated using the following formula:

$$Demand\ elasticity = \frac{\%\ change\ in\ quantity}{\%\ change\ in\ price}$$

To illustrate what this formula means in the context of health care we return to Jeff. We know from the previous example that with insurance, Jeff has to pay 10 Euros to visit the doctor. For the purpose of this example, let us say Jeff visits the GP five times a year. The next year, the insurer decides to raise the price of physician visits to 15 Euros (a 50% increase), and Jeff only visits the GP four times (a 20% decrease). This means that the elasticity of Jeff's demand for health care is -0.20 / 0.50 = -0.40. This number aids policymaking, as it helps predicting the changes in demand if deductibles or coinsurance rates would be increased or decreased. An intuitive way of interpreting the elasticity in this example is to look at it this way: If the price of a physician visit is doubled (100% increase), demand goes down by 40%.

Of course, policymakers are not interested in Jeff per se; rather, they wish to know what a population of Jeffs would do. So let us assume a demand elasticity of -0.40 and a population of a million Jeffs. With insurance, this population has five million physician visits every year (i.e., each member visits the physician five times a year). Increasing the price of a physician visit from 10 to 12 Euros (20%) would lead to -0.40 × 0.20 = 8% fewer visits. At the population level, this thus means 5,000,000 × 8% = 400,000 fewer physician visits.

Often, articles report *arc* demand elasticities where changes in both quantity and price are measured as percentages of the average of the old and new price/quantity. The elasticity we refer to above is a point estimate and reflects the elasticity of demand for health care given the current situation. This number is thus important for determining the effect of changes in a given health care system. Arc elasticities, however, refer to the elasticity of demand along a 'range' and might be seen as the best guess of the mean elasticity between two points on the demand curve. It is calculated as follows:

$$arc\ elasticity = \frac{\left(\frac{new\ demand - old\ demand}{(new\ demand + old\ demand)\ /\ 2}\right)}{\left(\frac{new\ price - old\ price}{(new\ price + old\ price)/\ 2}\right)}$$

This number is of interest when designing a completely new health care system, or when contemplating large changes in an existing health care system. Depending on the specifics of the situation, arc elasticities may be different from regular elasticities (e.g., arc elasticity of Jeff's demand for health care would be -0.55 instead of -0.40). However, the interpretation for both types of elasticity is similar: both give an estimate of how much demand decreases when price increases.

The RAND experiment

The RAND Health Insurance Experiment (Newhouse, 1996) is the pinnacle of experimental tests of moral hazard in the insurance industry and gives unique insights into the price responsiveness of the demand for health care. The Experiment was conducted between 1974 and 1982 and involved the random allocation of people in four large cities and two rural areas in the U.S. to different insurance plans. Approximately 2,750 families composed of 7,700 individuals participated for three to five years. Families were randomly assigned to different insurance plans so that researchers would be able to estimate the effects of insurance on the demand for medical care, without having to worry about selection problems. More specifically, each family was assigned to one of four fee-for-service plans with different levels of coinsurance (0%, 25%, 50%, and 95%; this is the percentage of the medical care costs that the insured pays for himself) or an HMO-style plan in which participants faced no cost-sharing and always needed a referral from their primary care physician to visit a specialist. The researchers also varied the maximum amount of out-of-pocket health care expenditures (5%, 10%, or 15% of annual income; the insurance covered all expenditures over this limit). The study tested how the different insurance plans affected the demand for hospital care, outpatient medical care, care related to acute and chronic episodes of sickness, dental care, and well-care (a label used for deferrable care such as immunizations, examinations, and birth control—deferrable does not mean unnecessary). The incredibly large amount of data gathered in this experiment has been described in more than 300 scientific articles and numerous reports. The interested reader is referred to a detailed book by the lead researchers in which they discuss the policy implications of the RAND experiment (Newhouse, 1996). In this section we highlight the most important findings with respect to how different levels of (co)insurance affect the demand for health care.

The main conclusion that may be drawn from the RAND experiment is that the demand for health care is sensitive to changes in price. The overall elasticity of demand for care health care is approximately -0.20. Yet, co-insurance and deductibles are crude and unrefined in the way they cut back health care expenditures. In addition, although the demand for health care decreases as the co-insurance rate increases, the largest decrease in demand is observed when co-insurance rises from 0% to 25%. This finding holds across almost all types of care, ranging from treatment of chronic diseases, to in- and outpatient care, dental care, and even to acute care such as emergency room visits.

There are some findings in the RAND experiment that require some elaboration. First, the researchers have gone to great lengths to test whether co-insurance increases *only* the demand for effective health care (something standard economic theory would predict). To this end, they had a panel of doctors go through a large portion of the medical files in the experiment to determine whether—given the disease or symptoms a particular patient had—the obtained care was appropriate or not. They found that co-insurance (across the different levels) decreases the demand for inappropriate care as well as for appropriate care. One might use this finding as an argument for free health care for all, to ensure that everyone gets the (appropriate) care they need. However, the authors are quick to point out that this finding may well be framed in opposite manner. Free care increases the demand for both inappropriate and appropriate care. Inappropriate care not only causes precious health care resources to go to waste; it may actually have negative health effects (side-effects of unnecessary drugs, complications that occur during superfluous operations, and so forth). The decreased use of health care is thus not unequivocally positive or negative.

Second, not all types of health care were equally sensitive to changes in price. More specifically, treatments that were deemed most easily deferrable—such as dental care and well care—showed the highest sensitivity to price (arc elasticities were -0.30 and -0.43 respectively). Historically, the economic argument has been to have greater coverage for treatments that are relatively less price elastic because they are less prone to perverse (ex post moral hazard) effects. The assumption thus is that the larger the elasticity, the higher the proportion of inappropriate treatments.

Looking at the price elasticities in the RAND experiment, one would then propose that dental and well care should have great levels of co-insurance, as their price elasticity is high. However, the relatively high level of price elasticity in these two categories is not due to a larger number of inappropriate treatments. Rather, people simply defer these treatments because they can. Again, the policy implications that follow directly from basic economic theory require some care.

Third, besides differences in elasticity between types of care, there is also considerable variance in estimates of elasticity for the same type of care along the range of cost-sharing possibilities. For example, between 0 and 25% co-insurance, the arc elasticity of demand for outpatient care is -0.17 while it is -0.31 between 25% and 95% co-insurance (Newhouse, 1993). This dissimilarity makes it difficult to use the estimates of the RAND experiment in health care systems that already have 45% cost-sharing, for example.

An important finding in the RAND experiment is that cost-sharing barely affects health outcomes for most people in the study. This combats the argument that people forego so much appropriate care under cost-sharing that it increases costs in the long run—through an increased need for more complicated and expensive care when untreated health problems deteriorate. Put differently, the free plan did also not markedly improve the health of its beneficiaries. This is likely due to the increase of inappropriate care on the free plan that—through side effects and complications—counteracted the beneficial effects of the additional appropriate care received. One important caveat is that those who were sick and poor at the beginning of the experiment did experience negative consequences from the cost-sharing plan. They exited the experiment with worse health than did the sick and poor on the free plan. This finding is corroborated by other studies that find that price elasticity of demand for medical care is highest for low-income people (Skriabikova et al., 2010).

Importantly, the long-term effects of being on the free plan or cost-sharing plan could not be tested because the maximum period of time people were on one of the experimental plans was five years. Indeed, a recent review of studies on the causal effect of health insurance on the demand for care and health outcomes (in U.S. samples) finds that insurance does improve health (Freeman, Kadiyala, Bell & Martin, 2008).

A final note on the RAND Experiment is that it did not test the ex post moral hazard hypothesis in the traditional sense: it did not compare the behavior of the insured to that of the uninsured. While the data are valuable for estimating the size of the effect of different levels of cost sharing, it is not clear what the demand for care of the uninsured would look like. However, it is clear from the data that there is no linear relationship between the level of cost sharing and demand for care. Therefore, extrapolating the RAND Experiment finding to no insurance is difficult, and estimating the net effect of some health insurance compared to no health insurance based on the RAND estimates remains tricky (but see Keeler, 1988, for an attempt to do so).

*A short note on demand elasticity***.** In the introduction of the section on ex post moral hazard we touched briefly upon the relative disagreement between different articles on the size of demand elasticity. A review of studies that estimate the elasticity of demand for physician services in U.S. finds that estimates are most often between -0.10 and -0.50 (Skriabikova et al., 2010). However, estimates range from smaller than -0.10 to greater than 1. A review that includes estimates from natural experiments as well as cross-sectional studies observes a similar disparity in the estimates and concludes that "the demand for health care falls with increases in out-of-pocket costs […] the magnitude of the estimated response varies widely, however" (Zweifel & Manning, 2000, p. 436). It is likely that the difference in the estimates between countries is due to differences between the specific health care systems. The best estimate of demand elasticity for a given country is thus one that is estimated using data from that same country. The next section reviews the (meager) empirical evidence on ex post moral hazard and the elasticity of demand for health care in the Dutch system.

Findings for the Dutch health insurance system

As noted, it is unclear to what extent the exact estimates of the elasticity of demand of one country or setting are applicable to another. Alarmingly, there are instances of considerably different estimates within the same country (see Cockx & Brasseur, 2003; and for a Belgian example, Van de Voorde, Van Doorslaer & Schokkaert, 2001). To get a sense of whether the Dutch system is subject to the same (moderate) ex post moral hazard effect that is found in other countries, we searched for ex post moral hazard tests in the Dutch system.

We could find only two empirical tests of ex post moral hazard in the Dutch health care system and both were conducted by the same author (Van Vliet, 2001, 2004). In the first paper, data on health care utilization, health insurance status, and background variables from 6,039 families (13,362 persons) who were privately insured between 1990 and 1994 were obtained from the Dutch Central Bureau of Statistics (CBS; Van Vliet, 2001). The author cleverly determined the average payment rate (similar to marginal cost of care; dependent on expectations of exceeding the deductible and health status and the size of the deductible) to estimate that the demand elasticity for health care on average is -0.08. There is almost no elasticity for hospital visits (-0.007) and relatively much for physiotherapy (-0.12). However, even the highest estimate is well below that of the RAND experiment. The author notes that this might be due to differences in the health care system between the U.S. and the Netherlands, and the relative size of health care spending in terms of GDP (9% in the Netherlands versus 14% in the U.S.; a higher share leads to higher price elasticity). Note that the sample under study is privately insured, and it is unclear whether the same results would hold for a publicly insured sample.

In a second article, Van Vliet (2004) estimates the elasticity of six types of care (general practitioner visits, physiotherapy, pharmaceutical care, medical specialist visits, hospital care and 'other') using data from a private health insurance company. For each of the 100,487 enrollees their expenditures are predicted if their deductible would be zero. Using the size of the deductible and the difference between the predicted and actual health care expenditures, the elasticity of the demand for each of the six types of care is estimated. Note that this procedure only allows estimating the elasticity of demand for the 48.7% of enrollees that have a deductible. The highest elasticity was found for GP visits (-0.40), followed by physiotherapy (-0.32) and the lowest for hospital visits (-0.04).

The same researcher—using two different datasets and two different methods—thus finds a rather low (-0.08 in 2001) as well as a more moderate (-0.14 in 2004) elasticity of demand for health care. In the 2004 paper the price elasticity for demand for visits to the general practitioner even is -0.4 while it is -0.09 in 2001. In both papers there is considerable variance in the demand elasticities across different types of care. The rank order of the estimates, however, is quite similar across the two papers.

These estimates leave the true demand elasticity of demand for care unknown. Yet, the findings would suggest that when policymakers aim to reduce health care costs, they should increase the cost of GP visits because in both papers the demand elasticity is highest for this type of care. However, physicians serve an important gatekeeper role in the Dutch health care system. Increasing the price of visits may lead people to circumvent the physician and visit a specialist right away. Specialist visits are generally more expensive, and reducing GP visits through increased cost sharing—a measure intended to cut costs—might thus lead ultimately to increased costs.

Policy implications of the findings on ex post moral hazard

It is clear from the findings reviewed above that ex post moral hazard exists. When the price of care goes up, demand generally goes down. However, we have seen that not all types of care are equally responsive to changes in price, and there are conflicting findings even for the same type of care (e.g., specialist visits). Worse, even within a health care system estimates of demand elasticity may differ wildly. While the difference seems small (the difference between estimates of elasticity for GP visits in the previous section is only -0.31), the effects on the population level are of huge importance. Namely, to decrease demand by a certain desired amount, low price elasticity would lead policymakers to increase deductibles or coinsurance rates substantially. In contrast, higher price elasticity would render a smaller increase sufficient. In a more extreme situation, low price elasticity may deter policymakers from introducing price increases at all, because they think that the additional burden placed on the insured does not outweigh the benefit of the slight decrease in demand.

Even though it is unclear whether the RAND experiment estimate is applicable to the Dutch situation, and whether it applies to the current (much more advanced) health care system, it seems that -0.20 is the least confounded estimate of the elasticity of the demand for care. We hasten to add, though, that the estimates for the elasticity of demand differ across the types of treatments. The treatments that were most easily deferred (without regard for their appropriateness) showed the highest price elasticity. In predicting the impact of certain policy changes, the RAND experiment estimates could thus be used, but with caution. At present, there is simply not enough evidence to determine with certainty the exact price elasticity of demand for care in the Dutch health care system.

A clear implication of the reviewed findings is that different levels of co-insurance may be desirable for different types of treatments. Again, however, determining the exact level of coinsurance remains tricky, as the available data do not render consistent estimates. In addition, a word of caution is warranted. Even if the true elasticity of demand were revealed, policymakers would still face tough decisions. The RAND experiment clearly demonstrated that increasing co-insurance or raising deductibles does not lead beneficiaries to select more appropriate care or cut inappropriate care. Co-insurance cuts the demand for care ‘across the board’. Ideally, insurers would distinguish between appropriate care and inappropriate care, and reimburse only the appropriate care while having 100% coinsurance for inappropriate care. However, the same treatment might be appropriate in one case but inappropriate in the other, and checking the appropriateness of each claimed treatment is simply not an option[7]. The insurance company thus needs data on which types of treatments are generally effective in which situation, in order to be able to adjust its coinsurance rates accordingly. This would require insurance companies to keep track of large amounts of data (e.g., health status, specifics of disease episode, treatment details, etc.), but this might prove worthwhile in the long run.

If insurance companies are interested in determining the exact elasticity of the demand for care, they should experiment a little bit. We have seen that tests of ex post moral hazard rely on datasets that have not specifically been created for the type of question that is asked. Often, proxies for health status (such as drug use, or number of physician visits in the previous year) have to be used and the distribution of insurance or different levels of coinsurance is endogenous. Insurance companies could boldly attempt a replication of the RAND Experiment, but this seems both expensive and impractical. Smaller scale experiments, however, seem more plausible and might prove very valuable. For example, with recent and upcoming increases in the mandatory deductible (Rijksoverheid, 2012) it is simple to offer a randomly selected (representative) sample of beneficiaries a lower deductible for free. The rejection rate for such an offer would be low and this set-up allows the experimenter to compare the demand for health care in the treatment group to a (randomly selected representative) control group that is not offered the same discount on its deductible. Monitoring these two groups for one year—measuring their health status and demand for care (both insured and uninsured)—could yield reliable estimates of demand elasticity of care in the Dutch health care system. Similarly simple designs could be created for specific treatments by making them cheaper or free for certain groups of randomly selected people. Controlled, randomized experiments like those just described could drastically improve knowledge of—and thereby predictions based on—the elasticity of demand for health care.

Insurance fraud

[7] Even determining the appropriateness of apparently meritless treatments such as homeopathy might prove difficult. While there is no scientific reason to assume that homeopathy works, it does work in the sense that it is accompanied by a placebo effect (Shang et al., 2005). Placebo effects may have real health benefits, but does this mean that they should be covered by basic insurance policy? We think not, especially given its relatively low price—people who want to use homeopathy are able to pay for their treatments themselves. However, there are divergent opinions on this matter, and those in charge of making policy should ultimately answer the question.

Checking every claim that is filed for truthfulness is highly cost-inefficient, and therefore many insurers approve claims having to trust that policyholders file claims *only* when they are legally entitled to do so. Economic theory predicts that the insured take the opportunity to betray the insurer's trust when they believe the benefit of filing an illegitimate claim outweighs the probability and costs of getting caught. Insurance fraud thus does not lead to greater use of insured services but it does lead to more or larger claims, thereby increasing the burden on the insurance system.

We have already established that fraud is an asymmetric information problem where the policyholder abuses the trust that the insurer is obliged to place in him or her. The ways in which policyholders can exploit their informational advantage are classified as either "hard" or "soft" fraud. Hard fraud includes deliberately staging accidents or faking claims to obtain a payment from the insurance company. Soft fraud includes 'claim padding' and opportunistically misrepresenting the facts on a claim to obtain a higher payment or a lower monthly premium (Insurance Information Institute, 2010). In Jeff's case, hard fraud means he fakes his medical bills, while soft fraud means he has a medical bill and changes some numbers to obtain a higher payment. The distinction between the two is not always as clear as in the previous example, but it seems that soft fraud is harder to detect than hard fraud.

Importantly, the available estimates of the prevalence of fraud are usually based on 'proven' instances of fraud (although it is not always clear how the estimates are derived). Therefore, it is doubtful that available estimates of insurance fraud capture all of the fraud that is committed. Nevertheless, the reports below indicate that insurance fraud is a significant problem.

The Insurance Information Institute (2010) estimates that 3 to 10 percent of U.S. health care expenditures are due to fraud. This translates to a loss of between $77 billion and $259 billion—obviously, a rather wide margin of error. The Coalition Against Insurance Fraud (CAIF) states on its website that $80 billion is lost annually to insurance fraud (across industries). However, they sum up health care fraud statistics from other sources that report estimates of 'at least 3 percent' of health care expenditures ($68 billion) as well as '$125 to $175 billion annually' (CAIF, 2012). The Dutch association of insurers (Verbond van Verzekeraars, 2011) reports the number of detected fraudulent claims (313 in 2010) instead of a percentage estimate. The 'Convenant aanpak verzekeringsfraude' (2011) reports that €7.5 million worth of fraud was detected in 2009. The Verbond van Verzekeraars (2011) reports a total volume of €84,524 million spent by health insurers. 0.000089% of health care expenditures would thus be due to fraud. These wildly differing assessments of fraud may reflect the effectiveness of the Dutch health care insurers (relative to U.S. insurers) in deterring fraud, a motivation on the part of U.S. insurers to make fraud seem as bad as possible, or a general lack of accurate information. Given that not all types of fraud are equally easily observable and—by its very nature—insurance fraud is concealed, we suspect that the latter is true.

While the literature on insurance fraud is growing, much of it focuses on the detection of insurance fraud and the deterrent effect of auditing (Derrig, 2002). Few articles examine the causes of insurance fraud or assume that asymmetric information and self-interest are the underlying causes. As a result, many articles are focused at particularities with respect to specific numbers or patterns within claims (e.g., Major & Riedinger, 2002) or present theoretical analyses of the effects of auditing (for an overview, see Picard, 2000).[8] We are not aware of any studies that systematically test the effects of punishment. However, a relatively straightforward prediction would be that greater punishment is associated with less fraud.

We believe that one implication of the fraud statistics reported above is that it might prove worthwhile to look at the causes of insurance fraud. A greater understanding of the motivation to commit fraud will lead to a better grip on how fraud can be prevented and might be more effective than punishing offenders after the fact—especially when it is so unclear how often fraud is actually committed. Therefore, we discuss the empirical literature on why people commit insurance fraud and discuss its policy implications.

Empirical evidence on reasons to commit insurance fraud

It is clear that the main reason for people to commit insurance fraud is that they want money. However, a focus on the financial gain could not explain why people commit insurance fraud per se; they might as well cheat on their taxes, steal a pair of shoes, or rig a boxing match they're betting on. Therefore, the evidence we discuss below centers around why people might feel that insurance fraud is justifiable. This way, we hope to shed light on the specific causes of insurance fraud to inform prevention strategies aimed at reducing fraud.

Before we discuss the scientific literature, we highlight a noteworthy investigation by the insurance industry. The CAIF (1997) held interviews in focus groups and surveyed a fairly representative sample of 602 U.S. respondents. Participants in the focus groups

[8] Auditing means that the insurer randomly checks the truthfulness of a subset of claims. This deters fraud because the insured cannot anticipate whether or not he will be audited. This economic analysis of the effects of fraud assumes, however, that people like Jeff decide to commit fraud solely based on the consciously calculated odds that their claim is audited—an unlikely proposition. Until 2012, there had not been a test where auditing probability was randomly varied in the field, and the accuracy of the mathematically predicted optimal auditing models is often not tested.

indicated that the most important reasons to consider filing a fraudulent claim were (1) saving money (2) getting unaffordable things done (3) getting back at insurance companies. The survey data revealed that 61% of respondents agreed that fraud is a way to seek “a fair return” on premiums. Furthermore, 56% assumed that insurance companies already charge a bit extra to cover the fraudulent expenses. These findings suggest that policyholders feel like insurance companies are making too much money over their backs and that fraud is a way of balancing the scales. Perhaps unsurprisingly, the report also finds that those with a more positive attitude towards the insurer were less tolerant of fraud.

The same report finds that people think the chances of being discovered are rather small, and approximately 60 percent of respondents indicated they thought that different types of fraud (e.g., claim-padding, misrepresenting an accident, or falsifying receipts) were “fairly” or “very” common (CAIF, 1997). Furthermore, almost one-third of the surveyed sample said they knew someone who had cheated on an insurance company but only 17 percent of those people had reported the cheater.

Interestingly, the survey also asked respondents to indicate what measures they thought would be effective in curtailing fraud. More than 90 percent indicated that claims should be verified more often—an expensive suggestion. Another much less expensive suggestion was to inform the public about the costs of fraud—we will return to this measure in the section on policy implications.

A survey commissioned by the Insurance Research Council focused on the relationship between the perception of insurance companies and perceived acceptability of insurance fraud (Tennyson, 1997). Specifically, a representative sample of 1,987 adults was asked how acceptable it is to increase a claim “to make up for the deductible” and “to make up for past premiums”. People who found paying the premium “very burdensome” thought insurance fraud was more acceptable than those who had less trouble paying for insurance. Also, policyholders who felt “very confident” about the financial stability of their insurer thought fraud was less acceptable than those who were less confident. These results were interpreted as evidence for a positive relationship between perceptions of insurance companies and perceived acceptability of fraud.

Another interesting finding from this article is that the tolerance of insurance fraud was strongly predicted by the average tolerance for fraud in the state where the respondent lived. This suggests that social norms are a strong predictor of how acceptable people think it is to cheat on the insurance company. Building on previous research, Tennyson (2002) analyzes another part of the data collected in the CAIF study reported above. The focus in this study is on how experience with insurance companies affects attitudes towards fraud. Specifically, those respondents who had filed an insurance claim in the two years before the survey were less accepting of fraud. Arguably, policyholders who get their claim approved have a more positive attitude towards the insurer. This finding is thus consistent with the claim that perceptions of insurance companies are correlated with attitude towards fraud.

The evidence discussed so far illustrates that policyholders often feel that it is not unethical to file a fraudulent claim and that the chances of getting caught are slim. However, these studies were run in U.S. samples and it is unclear whether the findings are applicable to the Dutch population. Luckily, the Dutch Coalition of Insurers—in collaboration with Leiden University—commissioned a survey that explored the determinants of insurance fraud among 50 (Dutch) policyholders who had committed fraud, and 51 who had not (Verschuur, 1992). The power of this survey is in the fact that the data are linked to information in the "Centraal Informatie Systeem" (CIS; a system that stores all Dutch insurance claims and uses algorithms to detect possibly fraudulent ones). This allowed the author to compare the answers of policyholders that are registered to have filed a fraudulent claim to those of policyholders who have never committed fraud.

The results of the survey are consistent with those of the U.S.-based studies. The two most important reasons to commit fraud were (1) financial need (2) making up for past premiums. Also, most respondents thought that the likelihood that a fraudulent claim would be discovered was small and the measure that was expected to lead to the greatest reduction in fraud was increasing the frequency with which claims are verified—again, an expensive undertaking for insurance companies.

The report by Verschuur (1992) also aims to get insight into whether fraudsters differ from non-fraudsters psychologically. On average, fraudsters are more sensitive to self-interest, and less to social norms and collective interests. They think insurance fraud is more acceptable (not surprising), are less satisfied with their socio-economic status, and are more competitive. Although this is interesting in itself, we think this finding does not help in fighting insurance fraud. It is not feasible to have every policyholder fill out questionnaires to determine a psychological profile—and consequently check the claims of those with a fraudster's profile more thoroughly. Additionally, it is likely that the people who would consider committing fraud, would also be likely to fake answers on such questionnaires.

Taken together, the empirical evidence indicates that policyholders think insurance fraud is common, acceptable, and easy to get away with. Importantly, it seems that most people feel the insurer-policyholder relationship is unfair, and that insurance fraud is a way to get back at evil insurance companies. In this light, the underlying cause of insurance fraud may become apparent.

It seems that policyholders do not construe insurance as a mechanism through which they share risk. Also, they do not consider the effects that fraud would have on the other policyholders of their insurance company. Rather, they feel like a big insurance company is undeservingly making money and that they do not get enough of a return on their premiums. This conceptualization of insurance is conducive to fraud because it makes filing illegitimate claims look like a victimless crime and a justifiable way to compensate oneself for the 'unfairness' of insurance contracts.

Indeed, Clarke (1990, p. 3) already mentioned that policyholders are likely to perceive insurance companies as "fair game". Policyholders are unaware of the fact that fraud eventually leads premiums to rise and that therefore, other policyholders bear the cost of the crime. Another point, also emphasized by Viaene and Dedene (2004), is that insurers should be careful not to assume that policyholders' knowledge of the workings and goals of insurance is correct. Given the responses in the surveys we reviewed, it is unlikely that people think of their insurance contracts as a means through which they transfer their risk of health care expenditures to the insurance company. Similarly, they seem unaware of the fact that the transfer is possible largely because the insurer pools the risks of many policyholders. In fact, it is ironic that people think fraud is a way to compensate for the money they 'lost' on premiums—not having to file an insurance claim obviously means that nothing bad has happened to them!

Policy implications of the findings on insurance fraud

Although the estimates of how much insurance fraud actually costs are only tentative, the evidence we reviewed revealed that insurance fraud is overwhelmingly regarded as common and acceptable. We think insurance fraud is the most wasteful form of moral hazard that we have described and it is worth fighting. We describe two lessons that can be learned from the reviewed literature and suggest ways to improve the way in which insurance fraud is combated.

First, policyholders should be educated about the nature of insurance. It is clear that most people lack an understanding of the essence of insurance—sharing risks—and this seems to lead to a somewhat tolerant attitude towards insurance fraud. Opportunities for education arise when new clients sign up or when policyholders file a claim. These are situations where the insurer is in close contact with its clients and allows for the dissemination of relevant information. The Inshared (Achmea) commercials are a good example of how policyholders may be informed. These advertisements show visually how a group of policyholders pools money that is redistributed to those who suffer misfortune. Such graphic illustrations of how insurance works are exactly the type of information that is needed. They may reduce the likelihood that people perceive insurance as an individual contract on which they are losing money. Instead, we expect, it leads them to realize that they are not 'alone' in the insurance contract and they get a sense of what happens with the money they spend on insurance.

Relatedly, it is important that insurers make sure policyholders realize that what they get for paying premiums is a reduction in risk. Knowledge that premiums are essentially the price of risk reduction will reduce the feeling that one's payments are in vain or that the policyholder is not getting his money's worth. Initiatives that give the policyholder insight into what 'his' money is used for can serve the same goal. When people are aware that their payments have been used to pay for someone else's care—or even better, that their own treatments are partly paid for by others—they will learn how insurance redistributes money to those that need it. Ultimately, such greater understanding should lead people to see the payments they receive from insurance companies as bonuses rather than a reduction of the amount the insurer still 'owes' them.

The second lesson is related to the widespread belief that insurance fraud often goes undetected. Although the articles we reviewed do not report a relationship between perceived auditing probability and attitude towards fraud, it is clear that the perceived frequency of successful auditing is extremely low. Most respondents therefore also indicated that they thought fraud would be most successfully reduced if insurers would check claims more often and more precisely. This is in line with theories that predict that greater likelihood of punishment deters people from filing illegitimate claims (for an overview, see Picard, 2000). The low perceived probability of getting caught would also reduce the deterrent effect of penalties that are laid upon fraudsters. If the perceived probability could be increased, the size of punishment could become an important factor in the decision whether or not to commit fraud—high penalties would have a greater deterrent effect than low penalties. However, we were unable to find empirical tests of this hypothesis.

The problem with checking claims is that it is expensive and the costs of the audit might be greater than the gain that results from rejecting a fraudulent claim. A relatively cheap way to increase the perceived likelihood of auditing among claimants is to explicitly mention that the claim they are filing gets checked using the CIS database. In addition, it might be wise to publicize cases (e.g., through press releases) in which fraudsters have successfully been caught—especially if the case is sensational. A long line of research in psychology teaches us that the things that come to mind easily are perceived to be more likely (e.g., Schwarz et al., 1991), and news items on detected insurance fraud will likely increase how often people think fraud is detected. Note that these suggestions can be implemented without having to actually check claims more often—and are thus relatively cheap!

Both implications—increasing awareness of the nature of insurance and increasing the perceived likelihood of fraud being detected—can be implemented credibly and successfully only when the message is perceived to be authentic. When the perception of insurance companies is negative, messages such as the ones we suggest may backfire. Namely, if policyholders would get the impression that the industries' attempt at reducing fraud is aimed at making more profit, they might become even more tolerant towards fraud. It is thus extremely important for insurers to make sure the image their policyholders have of them is positive.

A final comment on the fight against insurance fraud is that we need more data on actual fraud. While the surveys that measure tolerance of fraud are extremely valuable, the data on actual fraud are rather limited. It seems that the Dutch Coalition of Insurers has already started gathering more data as they distinguish between acceptance and claim fraud as of 2009. However, for each claim the CIS currently only saves name and address information, the branch in which the claim is filed, the amount claimed, and whether the claim is fraudulent or not. That renders it very hard to run informative analyses on the predictors of insurance fraud. We recommend that the CIS at least save what type of product or service is claimed and through which insurance company the claim is filed, as well as details with respect to the policy (such as level of deductible and how long the customer had been paying premiums). This would allow insurers to determine, for example, whether awareness-raising commercials such as those by InShared indeed lead to fewer instances of fraud. A comparison of amount claimed for the same type of product or service at two different insurance companies would also be informative. Namely—without having to look at cases where fraud is actually determined—higher average claims in one company could be indicative of claim padding (soft fraud). Similarly, comparisons of the extent to which insurance fraud occurs at different companies could teach insurers something about which fraud deterrence activities are successful and which are not.

Moral hazard in pensions

This section reviews the evidence on moral hazard problems in the pension industry. Specifically, for each of the three types of moral hazard we discuss to what extent they are applicable to pensions.

First, we provide a short description of the Dutch pension system. We return to Jeff, who could build up his pension through the three pillars that comprise the Dutch pension system. The first pillar guarantees that Jeff has a bit of income after he retires and is referred to as AOW (Algemene OuderdomsWet). Every taxpayer contributes to the AOW in a pay-as-you-go system where the current working population pays for the current retired population. The second pillar comprises an income-dependent contribution to a pension fund that is in charge of making sure Jeff's money is invested or saved, and returned to him after he retires. Participation in this second pillar is mandatory for many people who have a job (which is important for reasons we discuss later) and ensures that Jeff does not have to give up too much of his income once he retires. Jeff could also opt to add to his pension by using a third-pillar option such as investing in real estate or creating a stock portfolio.

To understand what ex ante moral hazard would look like in the Dutch pension system it is helpful to determine what risks people face and what preventive efforts they could engage in. Recall that ex ante moral hazard occurs when people take excessive risks or fail to engage in preventive efforts as a result of being insured. The risk is to be left with too little money to survive and live decently after retirement. The action to prevent this from happening is to save up enough money to pay for retirement. Without mandatory national pension schemes such as the Dutch system, people are completely responsible for their own pensions. In essence, by imposing the pay-as-you-go system, the Dutch government removes much of the responsibility for the pensions from individuals. Following the ex ante moral hazard argument, this should make people more careless about saving for retirement. The theory thus predicts that people would become more careless in planning their financial futures. The expectation that the government will take care of them once they retire might thus inhibit appropriate investing in private plans to supplement the public pension.

Although there are virtually no data that address this hypothesis, it is unlikely that people become less likely to save up for the future if the government imposes a mandatory pension system. Namely, it turns out that in situations where pension planning is not mandatory, participation is very low (e.g., Thaler & Bernatzi, 2004). A national public safety-net pension that guarantees a minimal standard of living could thus very well be enough to increase participation.

Following the ex ante moral hazard argument, however, the expectation that the government provides a pension would still decrease people's willingness to participate in the second pillar. After all, they know that a minimal standard of living is guaranteed—the responsibility for their financial future is partly transferred to the government—which would lead to a decrease in additional pension savings. As a result, retirees might have to give up a significant part of their living standard and possibly end up in financial trouble because they can no longer afford their mortgage.

Although it is analytically challenging to test for the kind of ex ante moral hazard suggested here, some successful attempts have been made. Most notably, Feldstein (1996) uses U.S. data on the generosity of the Social Security program and national saving rates to estimate to what extent public pension provision substitutes for private savings. Specifically, exploiting the variation in those numbers between 1930 and 1992, Feldstein (1996) finds that every dollar provided as Social Security wealth reduces private savings by two to three cents.

Returning to the Dutch pension system, it becomes apparent why strong encouragement of participation in the second pillar is important. Without mandatory participation in this pillar, the provision of AOW could decrease the likelihood that employees save up additional money for retirement. It thus seems that the Dutch system is well equipped to combat ex ante moral hazard. One might argue that the government should make participation in the second pillar mandatory to ensure that everyone is able to enjoy a comfortable retirement, but this question requires a political answer. In addition, as long as the current pension system provides a satisfactory solution, changes do not seem desirable.

To determine the extent to which ex post moral hazard poses a problem, it is again useful to determine what factors in the Dutch pensions system might be involved. Recall that ex post moral hazard occurs when a given service is used more often when it is insured than when it is not insured. The 'use' of retirement is non-flexible for everyone who lives long enough to exceed the retirement age. However, the possibility of early retirement does allow for flexibility in the age at which people retire. If the public pension provision is generous it might increase the prevalence of early retirement. So, without public pension provision Jeff might consider retiring at 65, or maybe even at 66 or 67, to obtain a little more financial room during his retirement. With the public pension provision, Jeff might not feel the need to save up a little extra money because he knows the government has saved up retirement money for him.

A review by Juurikalla (2008) summarizes the evidence that speaks to the question whether people retire sooner when the retirement benefits are more generous. All of the evidence points in the same direction—the generosity of public pension schemes and the generosity of their early retirement clauses positively predict early retirement rates. Put simply, Jeff is more likely to retire early if it means he gets generous retirement benefits, and less likely if the benefits are less generous. It should be noted, though, that this form of ex post moral hazard is not necessarily problematic in the same way ex post moral hazard in health insurance might be. In both cases people use the insured service more when it costs less. However, there is a case to be made that for health care decisions, there is an objective standard by which one could determine whether health care is needed or not. For health care decisions it is thus possible to label some treatments as 'inappropriate'. This is not (or is much less so) the case for early retirement. The fact that there is an option to retire early makes it appropriate to retire early. So, although people do seem to respond to the incentive effects of early retirement clauses, we are not sure that this behavior should be subsumed under the 'undesired' forms of moral hazard.

The final moral hazard—fraud—seems inapplicable to what most people think of when they think of pensions. Specifically, it is hard to fake retirement or to fake one's age to obtain retirement benefits earlier than legally allowed. However, an intriguing feature of many pension systems allows for some exploitation by its subscribers. Specifically, pension systems often have an optional disability insurance clause. Specifically, an enrollee in the pension fund can opt to pay an additional premium so that when he becomes (partially) disabled, the pension fund compensates the associated loss of income. Like any other insurance, this system may suffer from asymmetric information problems.

There is one specific disability insurance situation that goes to the core of asymmetric information: musculoskeletal disorders. Common examples include back, neck, and shoulder pains, as well as pain in joints and nerves. Such disorders are hard to diagnose, and it is difficult to objectively evaluate the seriousness of these problems, which allows for opportunistic behavior. So, imagine Jeff opted for the disability insurance in his pension plan and feels strain in his back. He goes to the doctor who determines that the back pain is severe enough to declare Jeff 40% disabled: because of the back pain he can only do 60% of the work that he used to be able to do. Consequently, the disability insurance clause in Jeff's pensions plan goes into effect and he starts receiving disability benefits to compensate for his loss of income. Usually, the doctor periodically checks the severity of the disorder that causes the partial disability. Under the disability insurance plan, Jeff has the incentive to overstate the pain he is feeling during the check-ups in order to keep receiving the disability benefits.

It turns out that exactly those diseases that are hard to diagnose—the musculoskeletal disorders—are responsive to the incentive effects in disability insurance. Variation in the generosity of two Canadian pension plans that provided disability insurance, and the rigor with which they checked claims based on hard-to-diagnose-disorders, has been linked to the prevalence of such claims (Campolieti, 2002). Specifically, the Quebec disability insurance plan was consistently tough on musculoskeletal-related claims between 1971 and 1999. However, the Canada plan increased the generosity of its musculoskeletal-related payments by $150 in 1987 and only adopted tougher medical screenings in 1995. Interestingly, the Quebec plan received only slightly increasing amounts of musculoskeletal-related claims over the entire period. In contrast, the Canada plan saw a much larger increase of musculoskeletal-related claims after 1987 and a sharp decrease after 1995. Notably, the claims related to diagnoses that are more easily diagnosed did not show the same systematic variation over time in either of the plans. These findings suggest that some claimants exploited the presence of asymmetric information, possibly by overstating the seriousness of their musculoskeletal disorders.

Moral hazard in pensions: conclusions

We have briefly discussed to what extent each of the three types of moral hazard applies to the Dutch pension system. In sum, the national public pension (AOW) could cause ex ante moral hazard by decreasing participation in pension systems that supplement the AOW. However, this could only happen if it weren't for the strong encouragement to enroll in these plans (participation is often mandatory).

The possibility for ex post moral hazard arises in the flexibility of the age at which people are allowed to retire. When early retirement benefits are generous, they encourage early retirement. To the extent that early retirement causes a loss of welfare (e.g., profitable labor is lost), it is worth considering whether early retirement benefits should be altered to discourage early retirement.

Pension providers usually also provide disability insurance plans. This makes them susceptible to fraud in that people may fake or exaggerate their disability to obtain disability benefits. Stringent medical screenings seem to effectively counteract this problem.

Discussion and conclusion

The empirical evidence for each of the three types of moral hazard is not equally strong. We have seen that there is very little evidence for ex ante moral hazard, although it may exist in cases where prevention of a disease is particularly hard and costly. The evidence for ex post moral hazard is much more consistent—almost all studies find that demand for care goes up when its price goes down. However, the estimates of the extent to which the demand for health care is sensitive to price are much less consistent. There is variance in the estimates of price elasticity for different types of care, but also for estimates of elasticity within types of care. The numbers on insurance fraud are by definition noisy estimates, as fraudulent actions are intended to go unnoticed but the available indices suggest that it is a significant problem. The main reason why people commit insurance fraud seems to be that they are unaware of the nature and goals of insurance.

To determine the extent to which moral hazard poses a welfare loss, it would be worthwhile to determine by how much moral hazard drives up insurance premiums. It would be even better to estimate how many people refrain from buying insurance who would have bought insurance, had the premium not been increased by moral hazard. In other words, does moral hazard lead to suboptimal levels of insurance purchased? We were unable to find literature that empirically estimates the size of this welfare loss, and we discuss below two reasons why we think it is unlikely that it will be precisely and reliably estimated in the near future.

First, the empirical challenge is substantial. Ideally, a researcher finds or builds a society and randomly offers one half of the people insurance while denying it to the other half. The researcher should control the way health care is provided and distributed to keep the true costs of health care constant. Also, the study should run for an extended period of time to allow both societies reach a stable level of demand for health care. This clearly is a daunting task; one that would dwarf the RAND experiment—in which people were 'merely' randomly assigned to different levels of insurance.

A second reason why it is unlikely that the welfare question will soon be answered is more theoretical. Specifically, it deals with whether economically optimal demand for care and optimal demand for insurance are equivalent to the desirable levels of demand for those services. We have seen that in the RAND experiment, the people with 95% co-insurance—which amounts to virtually no insurance—postponed or refrained from getting appropriate care. Although some might argue that it is economically optimal for people to deny health care if they do not value it at its cost, it is questionable whether such a situation is desirable. We already briefly mentioned that some ex post moral hazard might be construed as benign because insurance increases the likelihood that people get care when they need it. One could thus define the optimal level of demand for care as the level of demand in a society that does not have insurance, or as the level of demand—insurance-induced or not—that leads to the healthiest society. An answer to the question whether an insurance system converges to an optimal level of demand thus has a different answer depending on how the optimal level is defined. Even if the empirical challenge of estimating these optimal levels could be overcome, we are certain that researchers would disagree on which definition should be used.

A similar argument can be made with respect to the demand for insurance. People might decide not to buy insurance if it is too expensive, which could be considered economically optimal. However, again, one could also define the optimal level based on which level of insurance leads to the healthiest society. The estimated welfare loss—which would be estimated by comparing both observed levels of demand to their respective optimal levels of demand—would seem drastically different depending on which definition is used. In sum, although we think that the welfare question deserves an answer, the current empirical literature is unable to provide one.

We conclude this paper by reiterating a point that came up in all three discussions of policy implications: more data are needed. Currently, the data are not detailed enough to get a good understanding of how each type of moral hazard affects the Dutch health care system. Therefore, insurers are encouraged to collect more data on claim, health, and risk behaviors, as well as demographic information. In addition, insurers are advised to consider running experiments to get reliable estimates of how much of a problem each type of moral hazard is. The current climate—with budget cuts and changes in the health care system—is perfectly suited for randomly offering the possibility to some policyholders of keeping their more generous coverage while the coverage in the universal policy becomes less generous.

Overview of policy recommendations

1. Insurance companies need to collect more data to obtain insight into whether, when, why, and how people respond to insurance-induced incentives. Specifically, it is unclear which preventive health care behaviors are sensitive to ex ante moral hazard; the demand elasticity for health care is surely positive, but estimates vary considerably; and the true extent of insurance fraud is unknown. Collecting more detailed data is the first step that must be taken, but insurance companies are also encouraged to conduct simple experiments.
2. Ex ante moral hazard seems especially likely for preventive behaviors that are hard to maintain. Insurers should implement systems that aid policyholders in keeping up those behaviors—examples include encouraging exercising, enforcing chronic medication intake and using simple reminders (e-mail or text) to encourage patients to stick to a diet.
3. Policymakers are advised to use a demand elasticity of -0.20 when forecasting the effects of changes in insurance policies. They should be aware, however, that different levels of elasticity apply to different types of care and that all estimates have a wide margin of error. In addition, it is wrong to assume that cutting health care expenditures through cost-sharing leads to less demand for inappropriate care only, with equal or greater demand for appropriate care. Increased cost-sharing decreases the demand for both appropriate and inappropriate care.
4. Policyholders should be educated about the nature, goals, and workings of insurance. Insurance companies are seen as profiting from an unfair relationship with the policyholder, where the policyholder pays premiums but gets nothing in return. Specifically, insurers need to communicate that premiums are the price of risk-reduction, not advance payments for care. The money collected through premiums is mostly used for paying for (other) people's care, and thus does not disappear into the CEO's pockets. Additionally, the perceived likelihood of successful audits can be increased through publicizing detected cases of fraud.

Chapter 2 - Magical thinking in predictions of negative events: Evidence for tempting fate but not for a protection effect

Abstract. In this paper we test two hypotheses regarding magical thinking about the perceived likelihood of future events. The first is that people believe that those who "tempt fate" by failing to take necessary precautions are more likely to suffer negative outcomes. The second is the "protection effect", where reminding people of precautions they have taken leads them to see related risks as less likely. To this end, we describe the results from three attempted direct replications of a protection effect experiment reported in Tykocinski (2008) and two replications of a tempting fate experiment reported in Risen and Gilovich (2008) in which we add a test of the protection effect. We did not replicate the protection effect but did replicate the tempting fate effect.

This chapter is published as: Van Wolferen, J., Inbar, Y., & Zeelenberg, M. (2013). Magical thinking in predictions of negative events: Evidence for tempting fate but not for a protection effect. *Judgment and Decision Making, 8,* 45-54.

Students believe that they are especially likely to be called on to answer a question in class if they have not done the required reading (Risen & Gilovich, 2008), and people believe that they are especially likely to experience a mishap while traveling if they have not purchased travel insurance (Tykocinski, 2008). Both are instances of magical thinking where people who "tempt fate" by not taking necessary precautions feel that they are more likely to suffer negative consequences. Conversely, reminding people of precautions they have taken—for example, having purchased health insurance—leads them to see related risks as *less* likely (Tykocinski, 2008), a phenomenon we refer to as the "protection effect".[9] In the present research we examined both the tempting fate effect and the protection effect. We found consistent support for the tempting fate effect, but no support for the protection effect.

Tempting fate

When people tempt fate by neglecting to protect themselves from possible negative outcomes, they feel that those very negative outcomes are, ironically, more likely to occur. Risen and Gilovich (2008) detail how and why exactly the tempting fate effect occurs. Briefly, they argue that the act of tempting fate heightens the accessibility of negative outcomes. This heightened accessibility then leads to higher perceived probabilities of those outcomes (via the availability heuristic; Tversky & Kahneman, 1974). Tykocinski (2008) investigated how tempting fate beliefs affected the risk judgments of people who imagined having or not having insurance, and found that those who imagined that they were unable to purchase travel insurance believed that they were consequently at greater risk of losing luggage or needing medical care during their travels. Tykocinski interpreted this result as consistent with a belief in tempting fate: Failing to protect oneself by purchasing insurance brings negative outcomes to mind, which in turn makes those outcomes seem more likely.

Protection effect

In the research described above, Tykocinski (2008) also tested whether reminding people of precautions they have taken leads them to see associated risks as less likely. Specifically, she reminded people of their health insurance either before or after they rated the probability of needing medical care in the near future. Indeed, people who were reminded of their insurance before answering these questions thought they were less likely to need medical care than those who were reminded afterwards—the "protection effect". Tykocinski argued that this effect occurs because reminding people of precautions primes a general mindset of safety, making risks seem less likely.

[9] Throughout my dissertation I refer to this effect as the insurance effect, a term coined by Tykocinski (2013) after the publication of this paper.

The current research

While tempting fate and the protection effect might seem to be different sides of the same coin, there is reason to expect that the two effects might not be equally strong. Across many domains of judgment, “bad is stronger than good”—that is, negative information has stronger effects on judgment than does positive information (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Consequently, one might expect tempting fate beliefs, which are motivated by the heightened accessibility of negative outcomes, to show more robust effects on judgment than the protection effect, which putatively results from a mindset of safety. Here, we report five studies in which we examine both phenomena. We begin by reporting a study—which was conducted as part of a larger project concerning people’s thinking about insurance—in which we closely replicated the Tykocinski (2008) protection effect study described above. As we were unable to replicate the protection effect, we ran 2 additional replications in which we tried to stay as close as possible to the original study. These also failed to uncover any evidence for a protection effect. Finally, we report two conceptual replications in which we simultaneously tested both the protection effect and tempting fate. Here, we found evidence for tempting fate, but again found no evidence of a protection effect.

In each of the studies we report confirmatory analyses, in which we replicate the analytical strategy reported in the original papers. In personal communication, Tykocinski suggested that the protection effect would be more likely to occur for older people. To test such post-hoc explanations of failures to replicate, we report possible moderators such as age and gender in exploratory analyses sections where possible. In addition, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study (following the recommendations of Simmons, Nelson, & Simonsohn, 2012).

Study 1a: Tykocinski (2008) Exp. 1 with undergraduate subjects

This study aimed to replicate Experiment 1 reported in Tykocinski (2008). The hypothesis was that because, through commercials, insurance is associated with feelings of safety and protection, a reminder of insurance leads people to believe that they are less likely to be in need of medical care. We reproduced the procedure reported in Tykocinski (2008) as closely as possible, with the exception of the subject population. Whereas the original finding was based on data from train commuters in Israel, we ran the study in the Tilburg University social psychology lab and our subjects were Dutch undergraduate psychology students.

At the time we ran this study, we were not aware of the Simmons, Nelson, and Simonsohn (2011) paper that details how running unreported conditions and measures can lead to higher false-positive rates. While running the present experiment we ran four other conditions and included one extra risk-taking measure. In the procedure we describe below we report only the conditions in which we replicate the method reported in Tykocinski (2008) and leave out measures that were recorded *after* the original method. Since we do not find significant differences between conditions, higher false-positive rates are less of a concern. Nevertheless, a table describing the complete experimental design and measures is available in the Appendix.

Method

Subjects. Thirty-five Tilburg University undergraduate psychology students participated in a 60-minute research session of unrelated experiments that ran for a week in September 2010. They were assigned to a reminded (n=18) or non-reminded (n=17) condition. Gender and age were not recorded, but usually this group consists of 70% females around the age of 20.

Materials and procedure. The insurance reminder required people to indicate the name of their health insurance plan and whether they had additional coverage. Subjects then rated the extent to which they were satisfied with their health insurance on a scale ranging from 1 = “not at all satisfied” to 7 = “very satisfied”.

The reminder was either preceded or followed by 7 questions that required people to rate the probability of different events happening within the next five years on a 5-point scale (1 = “very small chance”, 5 = “very big chance”). Specifically, they rated the probability that during the next 5 years they would have to undergo a serious operation, would require physiotherapy, or would need to stay in the hospital for a long time. The original third question mentioned “comprehensive nursing care” (Tykocinski, 2008, p. 1348) but we changed the wording to make the question more easily understandable for undergraduates. The remaining four items required subjects to rate the probability that they would lose a substantial amount of money, that a war would break out in Europe, that they would win the lottery, and that Israel and Palestine would sign a peace treaty.

Results

All subjects indicated that they had health insurance. (This is unsurprising, as health insurance is legally required in the Netherlands.) Fifteen (42.9%) indicated that they had some form of additional coverage. Mean satisfaction level was 5.17 (SD = 1.01) and there was no significant difference in satisfaction level between the reminded (M = 5.22, SD = 1.12) and non-reminded condition (M = 5.12, SD = 0.93), $F(1, 33) = 0.09$, $p = .765$, $\eta^2 = .003$.

Confirmatory analysis. Mean evaluations of the probability of seven future events are shown in Table 1, along with univariate analyses of variance per item. The three health-related items were analyzed in a repeated-measures design with the reminder condition (reminded vs. non-reminded) as a between-subjects factor. Subjects who were reminded of their health insurance before they were asked about their likelihood of health problems did not differ in their ratings from those people who were reminded afterwards, $F(1, 33) = 0.053$, $p = .819$, $\eta^2 = .002$. In addition, after Bonferonni corrections, there were no significant differences on the four remaining measures.

The central result is that we did not replicate Experiment 1 in Tykocinski (2008). In retrospect we determined that this study was underpowered; using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) we found we had 76% power to find an effect as large as reported in the original study ($\eta^2 = .13$, cohen's $f = 0.39$). This might explain why we did not replicate the protection effect. In Study 1b, we ran a priori power calculations and determined that we needed at least 62 subjects to have 95% power to find an effect as large as reported in Tykocinski (2008).[10]

Another possible reason for this initial failure to replicate the protection effect was that our subjects were undergraduates whereas Tykocinski's subjects were commuters on a train. We were not in the position to fly to Israel to re-run the study with subjects from the original pool. We could, however, ask Dutch train commuters to fill out the survey. This is what we did in Study 1b.

[10] The original power calculations were done using G*Power 3.1. We estimated how many subjects we would need to obtain different levels of power to find a partial $\eta^2$of .13. G*Power 3.1 uses Cohen's f instead of partial η^2 but allows one to transform partial η^2 into Cohen's f, within the program. At the time of running these power analyses we did not know that there are multiple ways to compute partial η^2 and that one has to explicitly indicate what type of partial η^2 is used. Our original calculations assume G*Power's default-type partial η^2 while the actual partial η^2 was the SPSS-type. We redid the power calculations and found that our realized power values are lower than what we originally wrote. These are the actual realized power values for each study: 1a: 58.8% to find η^2 of .133; 1b: 84.7% to find η^2 of .133; 1c: 100% to find η^2 of .133 and 95.5% to find half that effect size; 2b: 92.5% power to find % η^2 of .049. [Note added Aug. 1, 2013]

Table 1. Means, standard deviations, mean ranks, sample sizes, and test statistics for all probability ratings of future events in Study 1a per condition.

	Reminded	Non-reminded	F	p	η^2
	M (*SD*)	*M* (*SD*)			
Operation	1.94 (0.80)	2.05 (0.83)	0.17	.681	.003
Physiotherapy	2.83 (1.20)	3.00 (1.06)	0.19	.667	.006
Nursing care	2.11 (1.18)	2.00 (0.70)	0.11	.740	.003
Monetary loss	2.28 (0.89)	2.35 (1.17)	0.05	.832	.001
War in Europe	1.67 (0.77)	2.42 (1.06)	5.70	.023	.147
Winning the lottery	1.28 (0.46)	1.06 (0.24)	3.31	.091	.084
Peace treaty	2.44 (0.51)	2.06 (0.97)	2.21	.146	.063
N	18	17			

Study 1b: Tykocinski (2008) Exp. 1 with train commuters

In the Netherlands, it is illegal to run studies in the train without a permit. Therefore, instead of in the train, commuters were asked to fill out the survey at or in front of the train station.

Method

Subjects. Seventy-eight commuters (M_{age}= 32.55, range 16–74; 42 female, 2 did not indicate gender) at the Tilburg Central train station voluntarily participated on December 20, 2012. They were randomly assigned to the reminded (n = 39) or non-reminded (n = 39) condition.

Materials and procedure. We used the same procedure as in Study 1a, but subjects received all instructions and questions in paper-and-pencil format. The insurance reminder required them to indicate the name of their health insurance plan and whether they had additional coverage. Subjects then rated the extent to which they were satisfied with their medical insurance on a scale ranging from 1 = "unsatisfied" to 5 = "very satisfied".

The reminder was either preceded or followed by seven questions. This time subjects answered exactly the same three health questions reported in Tykocinski (2008). Subjects rated the probability that, during the next five years, they would undergo a serious operation, would require physiotherapy, or would be in need of comprehensive nursing care. In addition, we asked them to rate the probability of two positive and two negative events. Specifically, subjects rated the probability that the current government would fall prematurely, that they would win the lottery within the next five years, that a European country would go bankrupt within five years, and that a Dutch person would win the Nobel Peace prize within five years. All questions were answered on scales ranging from 1 = "very small" to 5 = "very large".[11]

Results

Fifteen subjects did not indicate the name of their health insurers and were excluded from the analyses. Fifty (64.1%) indicated that they had some form of additional coverage. Mean satisfaction level was 3.84 (SD = 0.89) and there was no significant difference in satisfaction level between the reminded (M = 3.70, SD = 0.79) and non-reminded condition (M = 3.91, SD = 0.89), $F(1, 60) = 0.919$, $p = .342$, $\eta^2 = .015$.[12]

Confirmatory analyses. Mean evaluations of the probability of seven future events are shown in Table 2, along with univariate analyses of variance per item. The three health-related items were analyzed in a repeated-measures design with the reminder condition (reminded and non-reminded) as a between-subjects factor. Again, subjects who were reminded of their health insurance before they were asked about their likelihood of health problems did not differ in their ratings from those people who were reminded afterwards, $F(1, 60) = 2.75$, $p = .103$, $\eta^2 = .044$. If anything, the insurance reminder somewhat increased rather than decreased the probability ratings.[13]

[11] We thank Natascha Bauwens, Jolien Gordijn, Nienke Sterkens, and Maartje de Volder for collecting the data.

[12] Due to different amounts of missing data, the degrees of freedom vary among analyses.

[13] Including people who did not indicate the name of their health insurer did not meaningfully change the results, $F(1, 74) = 1.71$, $p = .195$, $\eta^2 = .023$.

Table 2. Means, standard deviations, sample sizes, and test statistics for all probability ratings of future events in Study 1b per condition.

	Reminded	Non-reminded	*F*	*p*	η^2
	M (*SD*)	*M* (*SD*)			
Surgery	1.79 (1.01)	1.42 (0.87)	2.69	.128	.038
Physiotherapy	3.03 (1.45)	2.75 (1.32)	0.62	.435	.010
Nursing care	1.34 (0.86)	1.09 (0.29)	2.57	.115	.041
Premature fall of government	3.45 (1.02)	2.94 (0.86)	4.52	.038	.070
Winning the lottery	1.31 (0.85)	1.30 (0.68)	0.001	.970	.000
European country bankrupt	3.62 (1.01)	3.30 (1.16)	1.30	.258	.021
Dutch Nobel Peace Prize	1.69 (0.76)	2.27 (0.91)	7.37	.009	.109
N	29	33			

Exploratory analyses. The age range of our subjects (16–74 years) is broader than that in Tykocinski's (2008) Experiment 1 (25–55 years) but it is possible that on average, she happened to recruit more older subjects than we did (mean age was not reported). If older individuals have stronger associations with insurance or are more concerned about negative health events, the protection effect might only occur in the older adults in our sample. To test this possibility, we included age as a covariate in the repeated-measures ANOVA and found that the probability ratings of the three events increased with age $F(1, 58) = 4.92$, $p = .030$, $\eta^2 = .061$. However, there was no effect of reminder condition, $F(1, 58) = 0.53$, $p = .473$, $\eta^2 = .009$. There was an almost-significant interaction effect between condition and age $F(1, 58) = 2.97$, $p = .090$, $\eta^2 = .049$. In a regression analysis where the reminded condition was coded as 1 and the non-reminded condition as 0, the coefficient for the interaction term (reminder x age) was positive but not significant for every item ($\beta_{surgery} = .505$, t = 1.75, $p = .086$, $\beta_{physiotherapy} = .257$, t = 0.87, $p = .388$, $\beta_{comprehensive\ nursing\ care} = .339$, t = 1.19, $p = .239$). The same analysis on a variable that is the sum of the three probability ratings paints a similar picture, $\beta = .485$, t = 1.72, $p = .090$. This indicates that, the older people were, the more likely probability ratings were to go up after the reminder. This is the opposite of the effect reported in Tykocinski (2008) Study 1.

We also tested whether the effect of being reminded of insurance on probability evaluations was different for men and women, but we found no main effect of gender, $F(1, 57)$

= 1.07, p = .306 η^2 = .018, and no gender x condition interaction, $F(1, 57) = 0.04$, p = .850, η^2 = .001.

In the current replication debate (e.g., Asendorpf et al., 2012), it has been suggested that variation in effects sizes may provide theoretical insights in the long run (IJzerman, Brandt, & van Wolferen, 2013). Therefore, we should run tests that have enough power to detect effect sizes that are smaller than the ones originally reported. In Study 1c, we determined that we needed 150 subjects to have 95% power to find an effect that was half the size (η^2 = .065, cohen's f = 0.26) of the originally reported effect size. However, if we would run 400 subjects we would have 95% power to find an effect with cohen's f = 0.16 ($\eta^2 \approx .025$) and 80% power to find an effect with cohen's f = 0.12 ($\eta^2 \approx .015$). So we decided to recruit 400 subjects on Amazon Mechanical Turk (MTURK).

Study 1c: Tykocinski (2008) Exp. 1 on MTURK

Subjects. Four hundred and three subjects completed the study on MTURK (M_{age}= 26.81, range 18–63; 136 female) in exchange for \$0.10 on December 3 and 4, 2012. People could only participate if they had an approval rate that was greater than 95% and if they lived in the U.S.[14]

Materials and procedure. We included an instructional manipulation check (IMC) to weed out inattentive subjects (see Oppenheimer, Meyvis, & Davidenko, 2009). Subjects were excluded from the study if they did not successfully pass the IMC. Five hundred and three people started the survey, 411 (81.71%) passed the IMC and 8 subjects did not finish, so we were left with 403 subjects with complete data.

We used the same procedure as in Study 1a and 1b. All instructions and questions were presented in subjects' web browsers using online survey software (Qualtrics). The insurance reminder required them to indicate the name of their health insurance plan and whether they had additional coverage. Subjects then rated the extent to which they were satisfied with their medical insurance on a scale ranging from 1 = "not satisfied at all" to 5 = "completely satisfied".

The reminder was either preceded or followed by seven questions. Subjects answered exactly the same three health questions reported in Tykocinski (2008), rating the probability that during the next five years they would undergo a serious operation, would require

[14] The "requester" on MTURK can approve or reject a "worker's" answers, so to obtain a 95% approval rate workers need to consistently deliver quality work. We restricted our sample to U.S. based subjects to obtain a somewhat homogenous group of subjects and to prevent people from developing countries—most likely without health insurance—from participating.

physiotherapy, or would be in need of comprehensive nursing care. In addition, they rated the probability that within next five years they would lose a large amount of money, that Europe would go to war, that they would win the lottery, and that Israel and Palestine would sign a peace treaty. All questions were answered on scales ranging from 1 = "almost zero" to 5 = "very high probability". Note that the questions and scale labels are exactly the same as reported in Tykocinski (2008).

Results

Unlike Israel or the Netherlands, not everyone in the U.S. has health insurance. Therefore, we coded whether subjects indicated the name of their health insurance companies. Forty-nine (12.16%) did not list a health insurance plan name or indicated that they had none. We exclude the people without health insurance from the analyses we report here, but the results are nearly the same when we include these people.

Fifty-six (13.90%) indicated that they had some form of additional coverage. Mean satisfaction level was 3.63 (SD = 0.92) and there was no significant difference in satisfaction level between the reminded (M = 3.63, SD = 0.89) and non-reminded condition (M = 3.61, SD = 0.95), $F(1, 352) = 0.07$, $p = .794$, $\eta^2 < .001$.

Confirmatory analysis. Mean evaluations of the probability of seven future events are shown in Table 3. The three health-related items were analyzed in a repeated-measures design with the reminder condition (reminded and non-reminded) as a between-subjects factor. Again, subjects who were reminded of their health insurance before they were asked about their likelihood of health problems did not differ in their ratings from those people who were reminded afterwards, $F(1, 352) < 0.01$, $p = .996$, $\eta^2 < .001$.[15]

Exploratory analyses. The size of this sample allowed a better test of whether the protection effect interacts with age, as suggested in Study 1b. We included age as a covariate in the repeated-measures ANOVA and found that the probability ratings of the three events increased with age $F(1, 350)$ 12.85, $p < .001$, $\eta^2 = .035$. However, there was no effect of reminder condition, $F(1, 350) = 0.93$, $p = .334$, $\eta^2 = .003$ and no age x condition interaction, $F(1, 350) = 0.98$, $p = .324$, $\eta^2 = .003$.

[15] Including people who indicated that they did not have health insurance did not meaningfully change these results, $F(1, 401) = 0.17$, $p = .717$, $\eta^2 < .001$.

Table 3. Means, standard deviations, sample sizes, and test statistics for all probability ratings of future events in Study 1c per condition for people with health insurance.

	Reminded	Non-reminded	*F*	*p*	η^2
	M (*SD*)	*M* (*SD*)			
Operation	1.74 (0.89)	1.75 (0.80)	0.005	.946	.000
Physiotherapy	1.71 (0.96)	1.71 (0.85)	0.005	.945	.000
Nursing care	1.34 (0.65)	1.34 (0.65)	0.001	.982	.000
Losing large sum of money	1.87 (1.03)	2.19 (1.05)	8.59	.004	.024
Europe goes to war	2.08 (1.05)	2.19 (0.88)	1.12	.291	.003
Winning the lottery	1.23 (0.69)	1.21 (0.67)	0.12	.730	.000
Israel-Palestine peace treaty	1.83 (1.00)	1.90 (0.85)	0.45	.504	.001
N	167	187			

We also tested whether the effect of being reminded of insurance on probability evaluations was different for men and women, but we did not find an interaction effect, $F(1, 350) = 0.11$, $p = .915$, $\eta^2 < .001$. There was a small main effect of gender: Women rated the three negative health events as slightly more likely, $F(1, 350) = 4.54$, $p = .034$, $\eta^2 = .013$.

The attentive reader will have noticed that we find some significant effects on the two positive and negative events that are not related to the health care. A reminder of health insurance led subjects in Study 1a to think war was less likely. In 1b, a premature fall of the government seemed more likely and a Dutch Nobel prize less likely after an insurance reminder. In 1c, subjects who were reminded of their insurance thought they were less likely to lose a large sum of money. Some of these apparent findings remain significant even after Bonferonni corrections. We believe these are examples of Type-1 errors but leave it to other researchers—who might have reason to believe these effects are real—to test whether they replicate.

In three separate studies, we were thus unable to replicate the protection effect reported in Experiment 1 by Tykocinski (2008). This failure to replicate was not due to insufficient power: In Study 1a, we had 76% power to find an effect as large as that reported by Tykocinski; in Study 1b, we had 95% power to detect such an effect, and in Study 1c we had 95% power to find an effect *half* the size of the originally reported effect. Our failure to replicate Tykocinski is also unlikely to be due to the use of undergraduate subjects, as Studies 1b and 1c used older subjects. However, the skeptical reader might feel that we are incapable of properly running experiments and that this explains our repeated failure to replicate the protection effect. (The first author readily admits that this thought crossed his mind as well.) In the following study we therefore attempted to test the tempting fate and protection effect hypotheses simultaneously. Specifically, we replicated the two "self" conditions of Experiment 2 reported in Risen and Gilovich (2008), which tests whether students believe that they are especially likely to be called on to answer a question in class if they have not done the required reading. We also added a condition in which we attempted to conceptually replicate the protection effect. In this condition, subjects were asked to imagine that they had prepared extraordinarily well. If, as the protection effect hypothesis holds, making precautions salient primes a feeling of safety that makes negative events seem less likely, subjects in this condition should think it *less* likely that they will be called on to answer a question.

Study 2a: Risen & Gilovich (2008) Exp. 2 + protection effect

Method

Subjects. One hundred thirty-three Fontys University at Tilburg students (93 female; M_{age}= 20.08; range = 17–28; 1 did not indicate age) participated in a 20-minute session of unrelated experiments that ran for 2 days in November 2011 in exchange for 4 Euros. They were assigned to either the "prepared" (n = 46), "did not prepare" (n = 42), or "prepared really well" (n = 45) conditions.

Materials and procedure. The experiment was programmed in Authorware 7.0 and subjects read on a computer screen that they were to imagine the following situation:

> *You are taking a course and you are in a work group with approximately 20 students. This work group weekly discusses a piece of text or an article. Everyone should read and understand the article prior to the work group at home.*

In the "prepared" condition subjects then read:

> *Like you do every week, you have read and understood the article reasonably well.*

The "did not prepare" condition was designed to make subjects feel like they were tempting fate and therefore they read:

> *You did not really have that much time this week so you chose to not do your homework once. You thus do not really know what the article is about.*

The "prepare really well" condition was designed to make subjects feel like had taken extra precautions and therefore they read:

You prepared really well this week. You read the article thoroughly twice and you even moved some appointments to make sure you had enough time to prepare for the lecture.

Subjects in all conditions then read the following:

This time, the teacher decides he will call on someone to publicly summarize the article in front of the group.

Subjects then rated the probability that the teacher would call upon them on a scale ranging from 1 = "very small chance" to 10 = "very large chance".[16]

Results

Confirmatory analysis. Subjects who imagined that they did not prepare for the lecture only thought it was slightly more likely (M = 6.19, SD = 2.04) that they would be called upon to publicly summarize the article than did those who imagined preparing (M = 5.24, SD = 1.84) or preparing especially well (M = 5.24, SD = 2.00), $F(2, 133) = 3.37$, p = .037, $\eta^2 = .049$. Post-hoc tests (LSD) indicated that the "did not prepare" condition differed from the other conditions ($p_{\text{prepare}} = .025$ and $p_{\text{prepare really well}} = .026$) whereas the "prepare" and "prepare really well" did not differ from each other, p = .990.

Exploratory analyses. In this study, we again tested for main- and interaction-effects of gender on the probability ratings but found neither, $F_{\text{main}}(1, 133) = 2.87$, p = .092, η^2 = .022, $F_{\text{interaction}}(2, 133) = 0.33$, p = .717, $\eta^2 = .005$. Perhaps because of a relatively restricted range of age, we do not find a very strong effect of age, $F(1, 132) = 3.07$, p = .082, $\eta^2 = .024$, or an interaction effect with age, $F(1, 132) = 1.57$, p = .212, $\eta^2 = .024$.

We thus replicated the tempting fate effect reported in Experiment 1 by Risen and Gilovich (2008). We added a condition in which people prepared especially well for the lecture to test whether this would lead to a protection effect. However, as one of the reviewers on a previous version of this article pointed out, we might not have given the protection effect a fair chance. Our control condition mentions preparation, while our protection effect condition mentions "preparing really well". The difference between these two conditions is not very large and a control condition that does not mention preparation at all might be better.

[16] Afterwards, we also asked subjects to indicate what the best preparation strategy would be in this case: 1 = do not prepare at all, 2 = prepare as usual, 3 = prepare really well. There is no difference between conditions in how this question was answered, $\chi^2(2, n = 133) = 0.59$, p = .75. No one indicated that one should not prepare and across conditions 82% indicated that one should prepare as usual, and 18% thought one should prepare especially well.

Therefore, in Study 2b we replicated Study 2a but altered the control condition so that it did not remind subjects of preparation at all.

Study 2b: Study 2a with a different control condition

Using G*Power we determined that we would need 251 people to find an effect as large as we did in Study 2a (η^2 = .049, Cohen's f = 0.23). To this end, we sent out the survey to 460 second-year undergraduate students—who had completed the third author's course 2 months earlier—on December 5th and closed the survey December 17 (although the last subject finished December 13). We also ran the study in the lab (which recruits from a different subject pool) between December 10 and December 14, 2012.

Method

Subjects. One hundred and eighty five people (40.2%) responded to the email and filled out the survey. One hundred and eighteen people participated in the lab; combining the lab and online responses yielded complete data for 292 people (M_{age}=20.6; range 18–36; 200 female). They were randomly assigned to the control (n = 97), tempting fate (n = 99), or the protection effect conditions (n = 96).

Materials and procedure. We included an instructional manipulation check (IMC) to weed out inattentive subjects (Oppenheimer, Meyvis, & Davidenko, 2009). Subjects could repeat the IMC if they failed to complete it successfully, but were automatically excluded from the study if they failed 4 times. However, every subject successfully passed before reaching the exclusion point.

The materials were identical to those of Study 2b with two exceptions. The survey was programmed in Qualtrics and we deleted the last sentence of the text that everyone read to ensure that the people in the control condition did not think about preparation for the class. The new text read:

> *You are taking a course and you are in a work group with approximately 20 students. This work group weekly discusses a piece of text or an article.*

In the control condition there was no additional text, while the other two conditions displayed the exact same text as in Study 2a.

We asked people to indicate where they filled out the survey (in the lab vs. at home or "other place") and at the end of the survey we asked people to indicate what the text they read said about preparation for class ("no preparation", "very good preparation", "no mention of preparation", or "don't know"). We included only people who passed this manipulation check (96.6%) and who took more than 10 seconds to read the text and answer the question (95.9%).

In total we excluded 20 subjects (6.8%) and ran the analyses on the complete data of 271 subjects.

Results

Confirmatory analysis. Subjects who imagined that they did not prepare for the lecture did not think it was more likely (M = 5.04, SD = 2.12) that they would be called upon to publicly summarize the article than did those who imagined preparing really well (M = 4.56, SD = 1.65) or who were not reminded of preparing at all (M = 4.71, SD = 2.10),F(2, 272) = 1.49, p = .227, η^2 = .011. Note that directional (one-tailed) tests of the tempting fate and protection effect support only the tempting fate effect, $t_{temptfate-protection}$ (171.99)= −1.74, p = .042, $t_{temptfate-control}$ (175) = −1.06, p = .144, $t_{control-protection}$ (159.42) = .52, p = .301. So, we find weaker evidence for the tempting fate effect than in Study 2a and find no support for the protection effect.

Exploratory analyses. We looked for main and interaction effects of age but found neither, F_{main}(1, 271) = 1.35, p = .589, η^2 = .004, $F_{\text{interaction}}$(2, 271) = 0.40, p = .674, η^2 = .003. Unexpectedly, we found a large difference in probability ratings between men and women, F(1, 272) = 27.66, p < .001, η^2 = .094, such that men thought they were less likely to be called upon (M = 3.85, SD = 2.13) than women did (M = 5.15, SD = 1.76). In addition, we found a significant interaction effect, F(2, 272) = 3.37, p = .036, η^2 = .025.

To explore this interaction effect we ran the ANOVA reported under “confirmatory analysis” above separately for men and women. For men, it is clear that there is no difference between the control (M = 3.86, SD = 2.33, n = 29), tempting fate (M = 3.56, SD = 1.95, n = 27), and protection effect conditions (M = 4.17, SD = 2.12, n = 24), F(2, 80) = 0.52, p = .598, η^2 = .013. Directional (one-tailed) tests do also not provide evidence for either effect, $t_{temptfate-protection}$ (49)= 1.07, p = .144, $t_{temptfate-control}$ (54) = .53, p = .299, $t_{control-protection}$ (51) = −.49, p = .312.

For women however, we replicate the findings reported in Risen and Gilovich (2008). Women who imagined that they did not prepare for the lecture thought it was more likely (M = 5.66, SD = 1.88, n = 65) that they would be called upon to publicly summarize the article than did those who imagined preparing really well (M = 4.69, SD = 1.46, n = 71) or who were not reminded of preparation at all (M = 5.14, SD = 1.84, n = 56), F(2, 192) = 5.38, p = .005, η^2 = .054. Post-hoc tests (LSD) indicated that the tempting fate condition differed from the protection effect condition (p = .001) but not from the control condition (p = .101). The control condition and the protection effect condition also did not differ from each other (p = .144).

Following these analyses we checked with the authors of the original tempting fate paper but unfortunately, age and gender were not recorded in the experiments reported in Risen and Gilovich (2008). We are not entirely certain what to make of this interaction-effect with gender. From personal experience with teaching female undergraduates we do think that they are more worried than male undergraduates about making public statements in front of class and we might have had insufficient power to detect this interaction in Study 2a. Following Risen and Gilovich (2007, 2008), people who can more easily imagine negative outcomes should be more likely to display a belief in tempting fate. The entirely post-hoc explanation that women can more easily imagine being embarrassed in front of class, and are therefore more susceptible to this specific demonstration of the tempting fate effect, is one that could be tested in future research.

Discussion

In three studies, we attempted to replicate the protection effect reported in Tykocinski's (2008) Experiment 1 as closely as possible. Using a student sample, a sample of train commuters (as in the original study), and a large online U.S. sample, we did not find evidence for this effect. In a follow-up study in which we tried to conceptually replicate the protection effect we also did not find support for it. However, we did find evidence supportive of a belief in tempting fate.

Why did the protection effect not replicate?

One possible reason for our failure to replicate Tykocinski's (2008) Experiment 1 is that risk may be a more salient factor in the daily lives of Israelis, compared to our Dutch and American subjects. Israel has a recent history of war, and today, bombings in public places and military conflict are still common. This might make Israelis more attuned to risk, and more sensitive to variations in it. If so, they might also be more susceptible to features of life that seemingly decrease the probability of misfortune (i.e., they are more sensitive to the protection effect). But note that there are also important similarities between those countries. Both Israel and the Netherlands require residents to purchase health insurance (and have done so for many years). In addition, before health insurance became compulsory in 1995 most Israelis also had health insurance (Israel Ministry of Foreign Affairs, 2002). Furthermore, in the U.S. sample (where health insurance is much less of a default than in Israel and the Netherlands), we also do not find evidence for the protection effect. Differences in how unusual health insurance is thus seem unlikely explanations for the differences in the findings we report and those in Tykocinski (2008).

There might of course be cultural differences in the extent to which people in different countries are susceptible to magical thinking effects in general (in this case, possibly because of a difference of risk-salience in the daily lives of the populations in question). However, we do find another form of magical thinking in Study 2a and 2b. This still leaves the possibility that the protection effect is more likely to happen in Israel than it is in the U.S. or the Netherlands, but that all populations are susceptible to tempting fate effects. This could be tested by simultaneously rerunning our Study 2b in Israel and the Netherlands.

A final possibility is that the protection effect reported in Tykocinski (2008) was merely due to chance. The conventional alpha levels do allow for 5% false-positives and it is possible that this study "accidentally" found a protection effect. The only real test of this possibility is to rerun the exact same study in Israel to see if the effect replicates.

Why attempt "direct" replications?

In light of recent discussions with respect to robustness of effects reported in the (social) psychological literature (Open Science Collaboration, 2012; Simmons et al., 2011) we feel it is important to point out that we did not just randomly pick one article to see if it replicates. We were (and are) genuinely interested in the protection effect as we thought that the insurance protection effect might be one of the causes of the moral hazard effect (i.e., insurance leads people to take more risk, Arrow, 1963). When we failed to replicate the original effect study reported in Tykocinski (2008) we tried harder to find evidence for the protection effect. As is clear from this paper, these efforts did not yield positive results.

Our attempt to replicate the tempting fate effect reported in Risen and Gilovich (2008) was aimed at testing whether we could find a different magical thinking effect. This would rule out the possibility that the Dutch are simply not sensitive to magical thinking effects. We thus think that the successful replication of the tempting fate effect adds credibility to the non-replication of the protection effect.

On a broader level, we think it is valuable to run direct replications to test the robustness and universality (i.e., cross-cultural robustness) of an effect. Initiatives like http://www.psycfiledrawer.org (see Carpenter, 2012) are a good start, but devoting some journal space to replication attempts seems valuable as well. In fact, many have argued that, without direct replication, scientific progress is difficult if not impossible (e.g., Feynmann & Leighton, 1997). In addition to direct replications, conceptual replications are important to test the generality of an effect and test its reliance on a specific method or paradigm (Nussbaum, 2012; IJzerman, et al., 2013). Here, of course, we report both: three direct replications (Studies 1a, 1b, and 1c) and two conceptual replications (Study 2a and 2b).

Finally, we stress that our failed replications do not necessarily mean that the protection effect reported in Tykocinski (2008) does not exist. We merely report that we cannot replicate this finding in the Netherlands, and that a conceptual replication also does not provide evidence for the existence of the protection effect. Future replication attempts will prove valuable, especially when aimed at detecting possible moderators that might explain our failure to replicate the protection effect.

Appendix

Full design of study 1a.

Condition	Order in which measures were administered		
1	**<u>Insurance reminder</u>**	**<u>Probability rating</u>**	Risk-attitude
2	**<u>Probability rating</u>**	**<u>Insurance reminder</u>**	
3	Probability rating	Risk-attitude	Insurance reminder
4	Risk-attitude	Insurance reminder	
5	Risk-attitude	Probability rating	Insurance reminder
6	Insurance reminder	Risk-attitude	Probability rating

Insurance reminder = 3 questions related to insurance described in Study 1a. Probability rating = the 7 probability ratings described in Study 1a. Risk attitude = Translated version of the domain-specific risk-attitude scale (Weber, Blais, & Betz, 2002). Measures in bold underlined font are reported in the paper.

Chapter 3 - Moral hazard: Not having insurance makes people careful

Abstract. We conduct three controlled tests of ex ante moral hazard and find that not having insurance makes people careful. Specifically, we assign people to an insured or uninsured condition, or to a control condition in which insurance is not salient. In two hypothetical situations (Study 1 and 2) we find that people in the uninsured conditions are less willing to take risk than people in the other two conditions. In Study 3 all participants play the same incentivized gamble but we label it as insured or uninsured, or we do not label it (control). We find that even in this case, people playing the uninsured gamble take less risk.

This chapter is based on: Van Wolferen, J., Van de Calseyde, P.P.F.M., Inbar, Y., & Zeelenberg, M. (2014) Moral hazard: Not having insurance makes people careful.

Insurance is the transfer of risk from a consumer or "policyholder" to an insurance company in exchange for payment. Buying insurance implies accepting a known small loss (in the form an insurance premium) in return for the compensation of possible larger losses when they occur. The insurance company has to set premiums low enough so that buying insurance is attractive to consumers, but high enough to make a profit. However, pricing policies is difficult because policyholders may take more risks once they have insurance. This is referred to as moral hazard[17] and it implies that insurance companies may not be able to set optimal prices if they base premiums on the risk level of the uninsured.

Moral hazard theory thus predicts that insured and uninsured people differ in the extent to which they take risks and/or preventive measures (Arrow, 1963; Pauly, 1968). Specifically, people with insurance are hypothesized to take more risks and fewer preventive measures compared to people without insurance, because insurance compensates for at least some of the negative outcomes associated with risk taking. Moral hazard implies that insurance cannot be priced accurately by observing the behavior (and thus the risk level) of uninsured people, because people act differently once they have bought insurance. Ultimately, moral hazard may lead to market failures, because increases in risk taking lead to premium increases, which may lead to cancellation of policies, which may lead to further price increases, which may cause even more people to cancel their policies, and so on.

If moral hazard does exist, insurers should worry because they might go out of business if they cannot price their policies right. Policymakers should also be concerned because implementing insurance reforms may have unintended effects on risk-taking and prevention, which could undermine the effectiveness of reforms. Empirical research, however, provides only limited support for the moral hazard hypothesis (for an overview of moral hazard in health insurance, see Van Wolferen, Inbar, & Zeelenberg, 2013b). For example, two studies found that obtaining health insurance led to more health risk-taking (Dave & Kaestner, 2009; Stanciole, 2008), but other studies have found that obtaining health insurance does not cause changes in risk prevention (e.g., Courbage & Coulon, 2004). Other studies report tests of moral hazard in very specific contexts (e.g., malaria-prevention in Ghana, Debebe, van Kempen, & de Hoop, 2012; diabetics' prevention efforts, Klick & Stratmann, 2007) that make the results difficult to generalize.

[17] Technically, in this article we are referring to *ex ante* moral hazard, not *ex post* moral hazard. For the sake of simplicity, we use "moral hazard" to refer to the differences in risk taking and prevention between people with and without insurance.

Furthermore, studies that test for moral hazard in insurance use field data and have to address adverse selection as an alternative explanation for the findings. Adverse selection entails that high-risk people are more likely to buy insurance than low-risk people and it predicts the same outcome as the moral hazard hypothesis: more risky behavior among the insured than the uninsured (Akerlof, 1970; Arrow, 1963; Pauly, 1968; Rothschild & Stiglitz, 1976). Often, studies exploit exogenous variation in insurance coverage to circumvent selection problems. However, this may lead to other selection problems, such as that there is only one particular subsample from the population available for the analyses (e.g., Dave & Kaestner, 2009, who tested the effect of health insurance in people who had never had insurance before turning 65 but received Medicare after turning 65).

The same difficulties apply to studies on moral hazard in car insurance where in some cases, selection problems lead to observed patterns inconsistent with moral hazard. For example, Chiappori and Salanie (2000) examined the effect of increased coverage on auto accident rates. Young drivers in France get a discount on their insurance premium if their parents are safe drivers. Those young drivers use that discount to buy greater coverage than their peers whose parents are not so safe drivers. Moral hazard predicts that greater coverage leads to greater risk-taking and thus more accidents, but the opposite pattern is apparent in the data. Perhaps because their parents are such great role models, the drivers with greater coverage actually have fewer accidents. In other cases, there are ways to at least partly circumvent selection effects (see, for example, Abbring, Chiappori, & Zavadil, 2008), but in general, drawing causal conclusions from correlational field data is extremely difficult

The aim of the present paper is to overcome the confounding factors, such as adverse selection, that are inherent in correlational data. Therefore, we test for moral hazard in three controlled experiments. Specifically, we examine the effect of insurance on people’s willingness to take hypothetical (Studies 1 and 2) and actual (Study 3) risks. In each experiment, we randomly assign people to one of three conditions. In the *insured* condition we explicitly tell people that they are insured, in the *uninsured* condition we explicitly tell people that they are uninsured, and in the *control* condition we provide no information about insurance status. We test the direction of the moral hazard effect by comparing the control condition to the insured and uninsured conditions. This allows us to examine whether being told that one is insured promotes more risk-taking, or whether being told that one is uninsured promotes carefulness (or, of course, whether both are true).

Moral hazard is a description of a putative empirical fact, not a psychological explanation of it. In fact, it may not require a psychological explanation at all. If people were rational utility-maximizers, they would be expected to take on more risk if they expect their losses to be defrayed by insurance. However, people are not rational utility-maximizers, at least not perfect ones. It is therefore worth considering psychological explanations for moral hazard as well. One such explanation is that thinking about having insurance gives people a sense of safety that makes them feel more sanguine about risks (e.g., Tykocinski, 2013, but see van Wolferen, Inbar, & Zeelenberg, 2013a). Another (complementary) explanation is that thinking about *not* having insurance makes people particularly likely to take precautions. This could happen because they see negative events as more likely to happen when they are uninsured, just as people think that it is more likely to rain if they leave their umbrella at home (Risen & Gilovich, 2008).

A final possibility is that people follow a heuristic of shunning any risk when uninsured. In any case, a strong effect of *not* having insurance on behavior would be consistent with findings from many domains that negative experiences and events have a larger impact on people's thoughts, feelings, and decisions than positive ones (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001).

The present research

In the first two studies we find that insurance status does affect risk-taking. Specifically, we find that people who imagine being uninsured are less willing to take risk than those who imagine being insured, or than those who do neither. We also examine several reasons why the insured and uninsured may differ in the extent to which they are willing to take risk. In Study 3 we build on these results and fully incentivize choices. In addition, in that study we eliminate the alternative explanation of rational responses to incentives by keeping the incentives participants face constant. The only difference between conditions is whether the situation is described as insured or uninsured and we find that even in this case, people who are told they are uninsured are less willing to take risk. We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Study 1: Skiing

In this study we randomly assigned people to a control, insured, and uninsured condition and asked them to indicate how much risk they would be willing to take. We tested which conditions differed from each other to test the direction of the moral hazard effect. In addition, we measured people's perceptions of the probability of misfortune (Tykocinski, 2008) to see if they might explain any differences in risk-taking.

Method

Ninety-one Tilburg University students[18] took part in an experimental session that was run for a week in December 2010, in exchange for course credit. They were randomly assigned to a control (n = 29) insured (n = 32), or uninsured (n = 30) condition. Participants were seated in individual cubicles where the experiment was run on a computer. All participants read the following scenario:

> *You are on a skiing trip. You are standing on top of the mountain and have time for 10 more rides downhill.*

In the control condition, no text was added. In the insured condition participants then read:

> *You have insurance for 'extreme sports'. This means that you are insured when you go out mountain biking, diving, or bungee jumping. You are also insured when you ski off-piste.*

In the uninsured condition participants read:

> *You do not have insurance for 'extreme sports'. This means that you are not insured when you go out mountain biking, diving, or bungee jumping. You are also not insured when you ski off-piste.*

All participants then answered the following three questions: "How many times (out of 10) would you go downhill via an off-piste route?" To measure how likely they thought misfortune was were asked, "how likely do you think it is something goes wrong if you go off-piste" on a scale that ranged from (1) 'very unlikely' to (10) 'very likely'. They also indicated their skiing ability on a scale that ranged from (1) 'cannot ski at all' to (10) 'I am very good at skiing'.

Results

Risk-taking (how often out of the 10 possible trips participants would go off-piste) differed between conditions, $F(2, 91) = 5.80$, $p = .004$, $\eta^2 = .12$. Tukey post-hoc comparisons revealed that there were no significant differences in risk-taking between the control condition (M = 5.59, SD = 2.89) and in the insurance condition (M = 6.00, SD = 2.72; p = .84). People in the uninsured condition (M = 3.67, SD = 2.95) did indicate that they would take fewer off-piste routes than people in the control condition (p = .03) and people in the insured condition (p <.01). Self-reported skiing ability did not differ between conditions (F < 1), but did have an effect on how often people said they would go off-piste, $\beta = .203$, $t = 1.96$, $p = .053$. Controlling for skiing ability does not alter the results. Finally, we found no differences in

[18] Due to a programming error, the demographic data were not saved properly, so we cannot report the exact data here. Usually, these student samples consist of 70-80 % females and the average age is around 20 years.

perceived likelihood of misfortune between the control (*M* = 5.34, *SD* = 2.77), insured (*M* = 5.22, *SD* = 2.30), and uninsured condition (*M* = 5.13, *SD* = 2.78), $F(2, 91) = 0.05$, $p = .952$, $\eta^2 = .001$.

Study 2: Window washing

We ran Study 2 for three reasons. First, we aimed to replicate the findings of Study 1 in a different context. Second, we wanted to test whether the moral hazard patterns we observed in the previous study were similar in cases where participants made risky decision for others rather than for themselves. Third and finally, we explored two plausible mechanisms that might cause the uninsured's decreased willingness to take risk. As in Study 1, we measured risk perception (Tykocinski, 2008) but we also measured anticipated regret, shame, and guilt. These emotions have been associated with inter- and intrapersonal harm (Zeelenberg & Breugelmans, 2008) and we thought they might explain differences between the self- and other-conditions in this study.

Method

Participants. We aimed to have 100 people per condition and five hundred and ninety-nine people (206 female, M_{age}= 29.16, range = [18-67]) completed the survey on Amazon Mechanical Turk (mTurk; Buhrmester, Kwang, & Gosling, 2011) in exchange for $0.20 on August 9, 2013. They could only participate if they had a 95% approval rating and were located in the United States, and they were only allowed to take the survey if they successfully passed an Instructional Manipulation Check (IMC; Oppenheimer, Meyvis, & Davidenko, 2009).

Materials and procedure. Participants were randomly assigned to one of the six conditions of a 2 (self vs. other) × 3 (control vs. insured vs. uninsured) design. In all conditions participants saw a picture of someone washing a window and read the following scenario:

> *For this study, please imagine that you and one other proprietor own a window-washing company. You and Tom, the other proprietor, are also the only employees of this company. Together, you take care of all the window-washing jobs your company accepts. You have been running this business together for the past 11 years.*
> *Today you received an unusual request from a company called APEZ. The window (one window) they ask you to wash is 20 feet up and they want it cleaned next Thursday. You and Tom usually only wash windows that are up to 15 feet up because of the risks involved with*

washing windows that are higher up. You usually charge $50 for a window that is 15 feet up.

In the ‘other’ conditions, participants then read:

> *You already have an appointment on Thursday, and you doubt whether you want to send Tom to do the job. On the one hand, you could charge extra because the job is different from what you usually do. In addition, it is legal for you to send Tom to do the job so you are not liable in case anything bad happens. On the other hand, working at 20 feet is risky for Tom. If he were to fall he could suffer severe injuries.*

In the ‘other-control’ condition (*n* = 102) no text was added. In the ‘other-insured’ condition (*n* = 75)[19], participants read:

> *You both have an excellent health insurance plan. The insurance policy also applies if you or Tom work at heights greater than 15 feet. So, if you were to send Tom to do the job, he would be working insured at a height of 20 feet.*

In the ‘other-uninsured’ condition (*n* = 100), participants read:

> *Normally, you both have an excellent health insurance plan. However, the insurance policy does not apply if you or Tom work at heights greater than 15 feet. So, if you were to send Tom to do the job, he would be working uninsured at a height of 20 feet.*

In the ‘self’ conditions, instead of the above texts, participants read:

> *Tom already has an appointment on Thursday, and you doubt whether you want to take the job. On the one hand, you could charge extra because the job is different from what you usually do. On the other hand, working at 20 feet is risky. If you were to fall you could suffer severe injuries.*

In the ‘self-control’ condition (*n* = 102) no text was added. In the ‘self-insured’ condition (*n* = 123), participants read:

> *You both have an excellent health insurance plan. The insurance policy also applies if you or Tom work at heights greater than 15*

[19] We planned to run 100 participants per condition per condition, but because assignment to conditions was truly random, cell size varied from 75 to 123.

feet. So, if you were to take the job, you would be working insured at a height of 20 feet.

In the 'self-uninsured' condition (n = 97), participants read:

Normally, you both have an excellent health insurance plan. However, the insurance policy does not apply if you or Tom work at heights greater than 15 feet. So, if you were to take the job, you would be working uninsured at a height of 20 feet.

All participants then indicated, "how much should APEZ pay for you to accept the job and (send Tom to) wash the window at 20 feet" on a scale that ranged from (1) '$50' to (7) '$350', which also had an option (8) 'I wouldn't take the job for any amount of money'.

In addition, we included two exploratory measures of whether the decision to accept the job was a moral or a financial dilemma. We anticipated that the decision whether or not to accept the job would be perceived more as a moral dilemma in the other-conditions, and more as a financial dilemma in the self-conditions. Therefore, all participants indicated to what extent they thought "deciding whether or not to take the job/send Tom was a moral dilemma" and the extent to which they thought "deciding whether or not to take the job/send Tom was a financial dilemma" on a scale that ranged from (1) 'not at all' to (7) 'very much'.

To measure risk perception we again measured the perceived likelihood of misfortune but this time we also asked about the perceived danger of the job. Risk is probability times outcome and this question was intended to capture the outcome-part of risk. All participants therefore answered the question "How dangerous is it for you/Tom to go up 20 feet and wash the window?" on a scale that ranged from (1) 'Not at all dangerous' to (7) 'Very dangerous'. All participants also answered, "How likely is it that you/Tom fall(s) from the ladder while working at a height of 20 feet?" on a scale that ranged from (1) 'Very unlikely' to (7) 'Very likely'.

Next, in three separate questions, all participants indicated how much "regret, shame, and guilt they would feel if they accepted the job/sent Tom to do the job and that you/he fell off the ladder while working at a height of 20 feet" on a scale that ranged from (1) 'Not at all' to (7) 'Very much'. Finally, they indicated their gender, age, WorkerID, and how they found the hit. They could also leave a comment if they wished.

Results

We analyze the main dependent variable (would participants accept the job, and if so, for how much money) using tobit, logistic, and OLS regressions. The results are provided in

Table 3. As reported below, we did not find any significant differences between the self- and other-conditions so we collapse across this factor in Table 3.

Table 3. Analysis of main dependent variable in Study 2

	Joint (Tobit)			Likelihood (Logistic)			Amount (OLS)		
	b (SE)	*z*	*p*	*b (SE)*	*Wald*	*p*	*b (SE)*	*t*	*p*
Intercept	3.65 (.17)	21.32	<.001	1.65 (.19)	74.79	<.001	2.78 (.09)	31.90	<.001
Insured	-0.40 (.24)	-1.63	.102	0.54 (.30)	3.19	.074	-0.08 (.12)	-0.66	.51
Uninsured	2.09 (.25)	8.27	<.001	-1.47 (.24)	38.28	<.001	0.54 (.14)	3.86	<.001
n	599			599			455		
Model test	-			$\chi^2(2)$ =76.62, $p < .001$			$F(2, 455) = 10.98$, $p <.001$		

Note: In Study 2, all participants indicated, "how much should APEZ pay for you to accept the job and (send Tom to) wash the window at 20 feet" on a scale that ranged from (1) '\$50' to (7) '\$350', which also had an option (8) 'I wouldn't take the job for any amount of money'. The Tobit regression analyses all data and treats '8' as the censoring point, taking into account the decision whether or not to accept the job as well as how much people charge *if* they accept the job. The logistic regression tests whether there are differences between conditions in whether people accept the job (1 = yes, 0 = no), without taking into account how much they charge. The linear regression only takes into account data from people who accept the job and tests whether the amount people charge differs between conditions. In each regression, two dummy variables test whether the insured and uninsured condition differ from the Control condition.

In the ‘self’ conditions, participants in the insured condition were slightly more likely to accept the job (94.3% did so) than people in the control condition (88.2%), and when they did they asked for approximately as much money (*M* = \$131.05, *SD* = \$47.39) as the people in the control condition did (*M* = \$141.65, *SD* = \$47.55). The people in the uninsured condition were less likely to accept the job (only 61.9% did so) and also asked more money if they did (*M* = \$165.84, *SD* = \$73.93).

In the ‘other’ conditions, participants in the insured condition were slightly more likely to accept the job (82.7% did so) than people in the control condition (79.4%), and when they did they asked for approximately as much money (*M* = \$142.74, *SD* = \$60.63) as the people in the control condition did (*M* = \$136.42, *SD* = \$57.02). The people in the uninsured condition were less likely to accept the job (only 47.0% did so) and also asked more money if they did (*M* = \$167.02, *SD* = \$66.17).

Using a MANOVA, we do not find any main or interaction effects on the perceived danger or likelihood of misfortune between conditions (all η_p^2’s <.01). We do find that anticipated regret, shame, and guilt are all higher in the ‘other’ compared to the ‘self’ conditions and we find a tiny effect of insurance-condition on the guilt ratings but no interactions for any of these variables. We also find that, on average, deciding whether or not to send Tom is more of a moral dilemma (*M* = 5.06, *SD* = 1.80) and less of a financial dilemma (*M* = 3.89, *SD* = 1.99) than deciding if participants would take the job themselves (moral: *M* = 2.95, *SD* = 1.74; financial: *M* = 4.99, *SD* = 1.64). In addition, we find an interaction such that this decision seen is most strongly seen as a moral dilemma in case Tom is uninsured. However, none of these patterns provides a parsimonious explanation for our main finding that people are more reluctant to take risk in the uninsured conditions. All ratings and tests are provided in Table 4.

Table 4. Means and standard deviations for all measures collected after the main dependent variable in Study 3, with F tests of main- and interaction-effects in the final column.

		Self	Other	
		M (SD)	M (SD)	MANOVA F-tests
Danger	Control	5.09 (1.36)	5.03 (1.21)	$F_{\text{Self-Other}}$ (1, 599) = 2.77, p =.097, η_p^2 = .005
	Insured	5.23 (1.12)	4.87 (1.46)	$F_{\text{Insurance}}$ (2, 599) = 2.71, p =.067, η_p^2 = .009
	Uninsured	5.38 (1.32)	5.26 (1.45)	$F_{\text{Interaction}}$ (2, 599) = 0.71, p =.493, η_p^2 = .002
Likelihood	Control	3.16 (1.26)	2.97 (1.32)	$F_{\text{Self-Other}}$ (1, 599) = 3.55, p =.060, η_p^2 = .006
	Insured	3.28 (1.30)	3.08 (1.49)	$F_{\text{Insurance}}$ (2, 599) = 1.73, p =.178, η_p^2 = .006
	Uninsured	3.44 (1.29)	3.19 (1.54)	$F_{\text{Interaction}}$ (2, 599) = 0.04, p =.966, η_p^2 < .001
Regret	Control	6.03 (1.46)	6.14 (1.36)	$F_{\text{Self-Other}}$ (1, 599) = 12.59, p <.001, η_p^2 = .021
	Insured	5.58 (1.67)	6.15 (1.36)	$F_{\text{Insurance}}$ (2, 599) = 2.77, p =.063, η_p^2 = .009
	Uninsured	5.91 (1.69)	6.51 (1.03)	$F_{\text{Interaction}}$ (2, 599) = 1.81, p =.165, η_p^2 = .006
Shame	Control	4.12 (1.95)	5.41 (1.71)	$F_{\text{Self-Other}}$ (1, 599) = 67.02, p <.001, η_p^2 = .102
	Insured	4.23 (1.91)	5.13 (1.90)	$F_{\text{Insurance}}$ (2, 599) = 2.01, p =.134, η_p^2 = .007
	Uninsured	4.26 (2.16)	5.82 (1.49)	$F_{\text{Interaction}}$ (2, 599) = 1.47, p =.232, η_p^2 = .005
Guilt	Control	4.29 (1.97)	5.85 (1.52)	$F_{\text{Self-Other}}$ (1, 599) = 119.90, p <.001, η_p^2 = .168
	Insured	4.19 (2.02)	5.84 (1.70)	$F_{\text{Insurance}}$ (2, 599) = 3.90, p =.021, η_p^2 = .013
	Uninsured	4.61 (2.28)	6.37 (1.27)	$F_{\text{Interaction}}$ (2, 599) = 0.15, p =.858, η_p^2 = .001
Moral	Control	3.12 (1.72)	4.66 (1.82)	$F_{\text{Self-Other}}$ (1, 599) = 208.64, p <.001, η_p^2 = .260
	Insured	2.90 (1.77)	4.87 (1.97)	$F_{\text{Insurance}}$ (2, 599) = 2.61, p =.074, η_p^2 = .009
	Uninsured	2.86 (1.73)	5.62 (1.50)	$F_{\text{Interaction}}$ (2, 599) = 6.40, p =.002, η_p^2 = .021
Financial	Control	5.03 (1.45)	4.04 (1.82)	$F_{\text{Self-Other}}$ (1, 599) = 56.86, p <.001, η_p^2 = .088
	Insured	4.74 (1.65)	3.85 (1.94)	$F_{\text{Insurance}}$ (2, 599) = 1.00, p =.369, η_p^2 = .003
	Uninsured	5.26 (1.76)	3.76 (2.19)	$F_{\text{Interaction}}$ (2, 599) = 1.60, p =.203, η_p^2 = .005

Study 3: Gambles

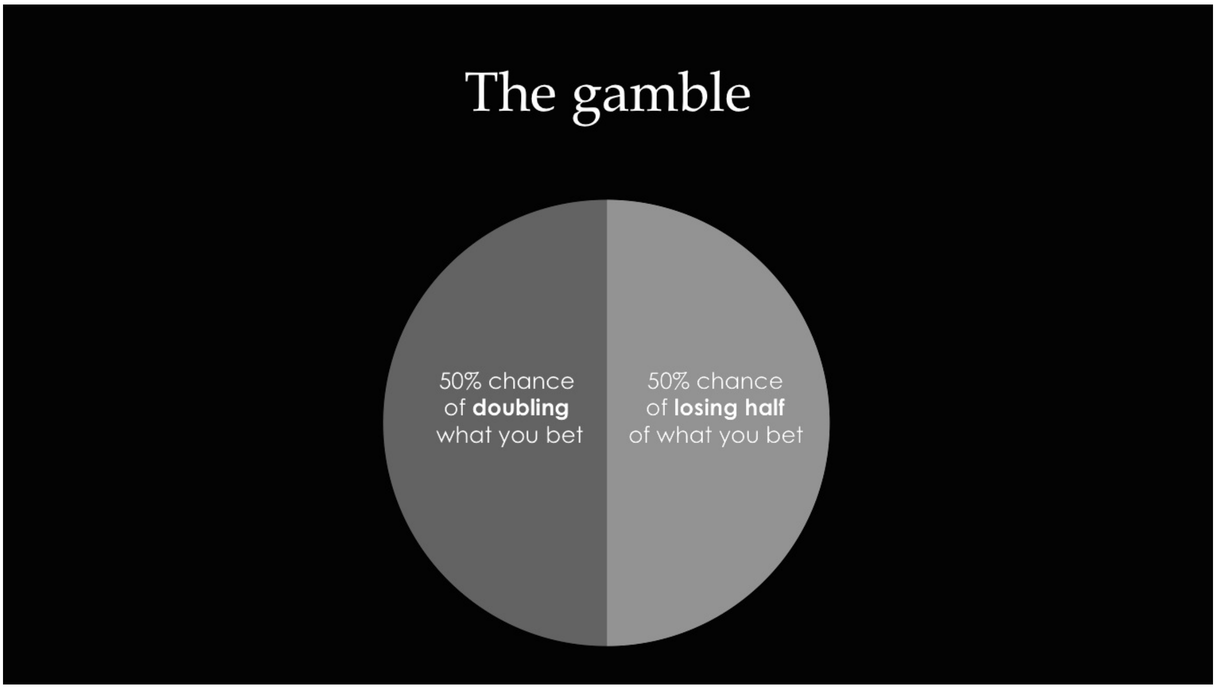
Although the rich context of the scenarios in our previous studies increased ecological validity, it also meant that we were unable to precisely equate the financial incentives participants imagined in the different conditions. Differences in willingness to take risk between those in the insured and uninsured conditions could therefore be the result of participants rationally responding to the fact that insurance would compensate them financially in the event of a negative outcome. In the current study, we examined whether simply labeling a situation as insured or uninsured affects people's risk preferences. Participants in every condition faced the same (incentivized) gamble—only the label changed across conditions. If we were to observe differences in risk-taking, these could only be due to the psychological effects of the insurance label, not to the financial incentives.

Previous research has tested how labeling options in terms of insurance affect risk preferences. In a choice between a small sure loss and some probability of suffering a greater loss, people are consistently more likely to choose the sure loss if it is described as 'paying insurance premium' than when it is described as a sure loss (Hershey & Schoemaker, 1980; Schoemaker & Kunreuther, 1979; Slovic, Fischhoff, & Lichtenstein, 1988). Similarly, in one study probability weighting functions were estimated based on more than 600 binary decisions between a sure loss and a gamble that each participant made (Kusev, van Schaik, Ayton, Dent, & Chater, 2009). From the choices in these decisions it could be inferred that the overweighting of small probabilities was greater when the sure loss was presented as an insurance premium than when it was presented as a sure loss. This context-dependence of risky choice has been explained in several ways (e.g., change of reference-point, increased salience of social norms to be prudent, over-weighting of small stated probabilities in insurance-contexts). However, none of the explanations make specific predictions about differences in risk-taking depending on whether the situation is labeled as insured or uninsured. Our study is the first to explore this issue.

Participants. We aimed to have 100 people in each condition. Three hundred and seven people (154 female, M_{age}= 33.33, range = [19-66]) completed the survey on mTurk in exchange for $0.10 between January 14 and January 17, 2014. They could only participate if they had a 95% approval rating and were located in the United States, and they were only allowed to take the survey if they successfully passed an attention check (Oppenheimer et al., 2009). In addition, participants received $0.10 bonus that they could use to bet on the gamble in the study.

Materials and procedure. Participants were randomly divided between three conditions (control vs. insured vs. uninsured). In all conditions participants learned that they would be playing a gamble and that they would receive a $0.10 bonus that they could use to play the gamble.[20] The gamble consisted of a 50% chance to double what they bet, and a 50% chance to lose half of what they bet (see Figure 1). Participants could choose to bet 0, 2, 4, 6, 8, or 10 cents.

Figure 1. The gamble all participants played in Study 3.



Note: All participants could choose to bet 0, 2, 4, 6, 8, or 10 cents on this gamble.

[20] $0.10 might seem small but it actually is a decent bonus to mTurkers who work for $0.10 per minute ($6 per hour). The bonus could more than double their usual pay.

In the control condition participants then read two examples of how the bet might turn out and chose how much to bet (for a full description of the materials and procedures see the Appendix). In both insurance conditions participants learned that they would be playing with or without insurance, and subsequently, whether or not they were insured. Even though every participant faced the same gamble, we made it seem plausible that there was a difference between the insured and uninsured participants. Specifically, in the insured condition participants read:

> *You can choose to bet 0, 2, 4, 6, 8, or 10 cents. The gamble gives you a 50% chance of doubling what you bet and, since you are insured, a 50% chance of losing half of what you bet.*
>
> *If you had not had insurance, you would have lost everything in case you gambled and lost.*

In the uninsured condition, participants received the following information

> *You can choose to bet 0, 2, 4, 6, 8, or 10 cents. The gamble gives you a 50% chance of doubling what you bet and, since you are uninsured, a 50% chance of losing half of what you bet.*
>
> *If you had insurance, you would not have lost anything in case you gambled and lost.*

Like the participants in the control condition, participants in both insurance conditions read two examples of how the bet might turn out and then chose how much to bet. After placing their bet, all participants learned the outcome of their bet and were told that they would be paid accordingly within 7 days. They then completed a final attention check and provided demographic information.

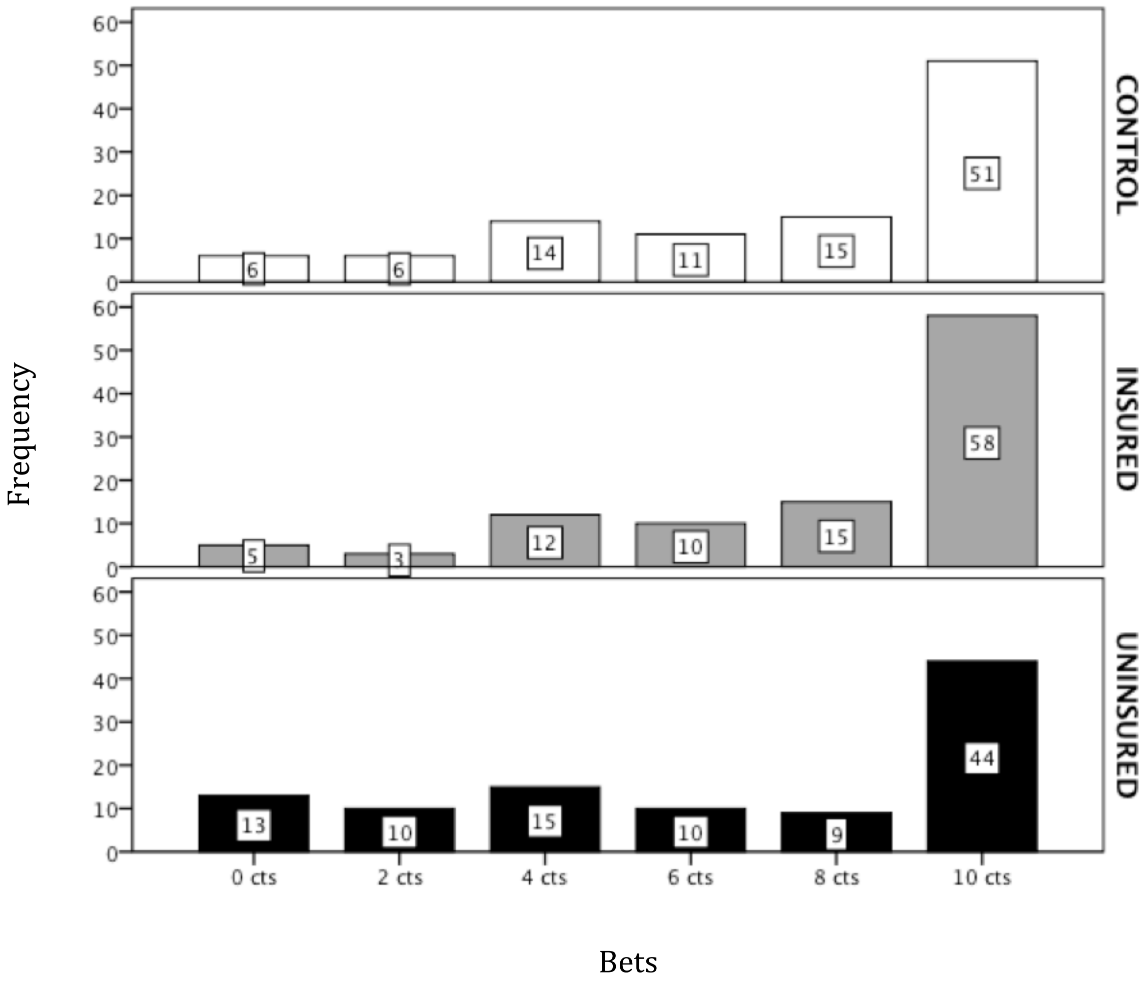
Results

Using an ANOVA, we find that people bet approximately the same amount in the control (M = 7.42, SD = 3.19, n = 103) and insured condition (M = 7.90, SD = 2.96, n = 103), while people in the uninsured condition bet less (M = 6.46, SD = 3.75, n = 101), $F(2, 304) = 5.04$, $p = .007$, $\eta^2 = .032$.[21] In LSD post hoc comparisons, the uninsured condition differs from both the control (p = .039) and insured condition (p = .002). See Figure 2 for a plot of the data.

[21] A non-parametric Kruskal-Wallis test, corrected for ties, is also significant, $\chi^2(2, N = 307) = 7.46$, $p = .024$.

We analyzed performance on the (final) attention check and found it did not have a main effect on how much people bet ($F(1, 301) = 0.22$, $p = .643$, $\eta^2 = .001$, nor did it interact with condition $F(2, 301) = 1.30$, $p = .276$, $\eta^2 = .009$. The effect of how the gamble was labeled remained significant when we controlled for performance on the attention-check $F(2, 301) = 3.81$, $p = .023$, $\eta^2 = .025$.

Figure 2. Distribution of bets per condition in Study 3. The numbers indicate how many participants bet that amount in that condition.



Discussion

In three studies we tested the existence and direction of the moral hazard effect by comparing a control condition in which insurance is not mentioned to an insured and uninsured condition. We find that people are less willing to take risks when they are uninsured, compared to when they are insured or when insurance is not mentioned. We found this effect in a hypothetical decision to take physical risk (Study 1 & 2), in a hypothetical decision to accept an assignment that involved physical risk for someone else (Study 2), and in an incentivized decision that involved financial risk (Study 3).

One of the strengths of our approach is that we consistently observe the same pattern of differences in risk-taking between conditions across three very different domains. More specifically, even though the surface structure of our experiments is very different (skiing, washing windows, playing a gamble), we obtain consistent results between them. This suggests that we successfully addressed the same "deep structure" in all three experiments, namely that of whether and how risk taking decisions are affected by insurance status (Wagenaar, Keren, & Lichtenstein, 1988).

We examined two possible mechanisms that might lead people without insurance to reduce the amount of risk that they are willing to take. In Study 2, we measured the anticipated negative emotions mainly because we expected they could explain possible differences between the self- and other-conditions. However, in both conditions we find that the uninsured are less willing to accept the job and that they ask more money *if* they accept the job. In addition, we find that differences in anticipated shame, guilt, or regret could not parsimoniously explain differences between the three insurance conditions.

We also measured perceived risk in Study 1 and 2 and we observe the risk *taking* patterns that we predicted based on the literature on risk *perception* (Risen & Gilovich, 2008; Tykocinski, 2008). Inconsistent with that literature however, we did not find that people in the uninsured conditions think risk is more likely. Perhaps this is because we measured risk perception rather crudely on 7- or 10-point scales. Although this is consistent with the literature on risk perception, these scales have been found to underperform in accurately measuring perceived risk (Haase, Renkewitz, & Betsch, 2013). Maybe unsurprisingly then, we also only find weak correlations between willingness to go off-piste and perceived risk in Study 1 (r = -.21, p = .050), and between amount of money requested to accept the job and perceived danger (r = .15, p = .002) and likelihood of misfortune (r = .11, p = .021) in Study 2. It might thus be the case that our measures of risk perception were not sensitive enough to pick up differences between conditions. In future work, researchers could consider using alternative measures to more accurately assess perceived risk, but currently we cannot support claims that differences in risk perception may partly explain differences in risk taking.

There may be a simpler explanation for our findings: People heuristically decide that they should not take risk if they do not have insurance. This would fit within a reason-based choice perspective that states that when people need to choose, they find or come up with reasons and justifications for their decision (Shafir, Simonson, & Tversky, 1993). When people face uncertainty without insurance they have a clear reason not to take risk if they do not have to: they are fully aware that if things go awry, they will suffer the consequences of their actions. Perhaps, not having insurance leads people to consider the possible negative outcomes of their actions and to overweigh them when deciding what to do. Conversely, having insurance does not provide participants with a very strong reason to either take or not take risk. It does not highlight the possibility of misfortune nor does it put a spotlight on the possible benefits of risk-taking. Having insurance may give people peace of mind, but it has yet to be established that this leads to a greater willingness to take risk.

The findings in this paper also relate to literature on how the salience of incentives affects decisions. People have limited attention and cannot always attend to all the features of the situation that may affect their outcomes (see section 4.2 in DellaVigna, 2009). A prime example of how incentives affect decisions more when they are salient is that alcohol sales reduce most if tax increases are included in posted prices, rather than at the register (Chetty, Looney, & Kroft, 2009). In our Studies 1 and 2, it is very salient to participants that they are insured or uninsured; we explicitly state that they are when we describe the situation they have to imagine themselves in. If the effect we obtain in those studies were only due to incentives, the salience of the insurance status might have led to an overestimation of the size of the effect. In Study 3 however, we led participants to believe that they were insured or not while in reality everyone faced the same incentives. The fact that, even in this case, we find that the uninsured are less willing to take risk suggests that our findings are not merely the result of the salience of the incentives people faced.

Before closing, let us return to the question that spurred the current studies. We believe that we have provided clear experimental evidence for the existence of moral hazard. This is important because although studies that use field data have great value, they suffer from selection- and other problems that may not allow for tests that reveal the true nature of moral hazard. In our studies, moral hazard seems to stem from people who are uninsured being more careful than people who are insured. This is important for policy makers and insurers because it suggests the largest change in risk behavior may be observed after reducing coverage, rather than after expanding coverage.

Appendix: complete description of Study 3

Here we report the details that we left out of the method section of Study 3.

Method

Participants were randomly divided between three conditions (control vs. insured vs. uninsured). In all conditions participants learned that they would be playing a gamble and that they would receive a 10-cent bonus that they could use to play the gamble:

> *You will receive 10 cents for completing this HIT. In addition, you receive a 10 cent bonus with which you can do whatever you want in this study. You can choose to keep the 10 cent bonus or you can choose to use it to bet on the gamble that we explain on the following pages. In any case, please know that you will receive 10 cents for completion + whatever your bonus is at the end of the study.*

In the control condition, participants then saw the gamble they were playing (Figure 1) and they read the following instructions.

> *You can choose to bet 0, 2, 4, 6, 8, or 10 cents. The gamble gives you a 50% chance of doubling what you bet and a 50% chance of losing half of what you bet.*

On the next page, participants read the following instructions and they could only proceed if they indicated they understood how the gamble worked.

> *Here are two examples of how the bet may work out.*
>
> *Example 1: if you choose to bet 6 and you <u>win</u>, you will receive:*
>
> ***10 cents** for completing the HIT*
>
> ***4 cents** that you did not use to bet*
>
> *6 cents that you bet x 2 = **12 cents** as your bonus*
>
> ***total = 26 cents***
>
> *Example 2: if you choose to bet 6 and you <u>lose</u>, you will receive:*
>
> ***10 cents** for completing the HIT*
>
> ***4 cents** that you did not use to bet*
>
> *6 cents that you bet divided by 2 = **3 cents** as your bonus*
>
> ***total = 17 cents***

On the next page, participants indicated how much they wanted to bet.

In the insured and uninsured condition, participants played the exact same gamble but we labeled it as insured or uninsured. In both these conditions, participants first read:

> *In this study, some people will play the gamble with insurance while others will play the gamble without insurance. The computer will randomly determine whether you are playing the gamble insured or uninsured. Please click the button below to find out whether you are insured or uninsured!*

In the insured condition, participants then learned that they were insured and proceeded to the gamble. They received the exact same information as the people in the control condition, but we added text that made it plausible they were playing an insured gamble:

> *You can choose to bet 0, 2, 4, 6, 8, or 10 cents. The gamble gives you a 50% chance of doubling what you bet and, since you are insured, a 50% chance of losing half of what you bet.*
>
> *If you had not had insurance, you would have lost everything in case you gambled and lost.*

They then received these instructions and could only proceed if they indicated they understood the gamble:

> *Here are two examples of how the bet may work out.*
>
> *Example 1: if you choose to bet 6 and you <u>win</u>, you will receive:*
> ***10 cents*** *for completing the HIT*
> ***4 cents*** *that you did not use to bet*
> *6 cents that you bet x 2 =* ***12 cents*** *as your bonus*
> ***total = 26 cents***
>
> *Example 2: if you choose to bet 6 and you <u>lose</u>, you will receive:*
> ***10 cents*** *for completing the HIT*
> ***4 cents*** *that you did not use to bet*
> *You are insured so:*
> *6 cents that you bet divided by 2 =* ***3 cents*** *as your bonus*
> ***total = 17 cents***
> *Remember that if you had not had insurance you would have lost all of what you had bet (last sentence in red).*

In the uninsured condition, participants learned that they were uninsured and read:

You can choose to bet 0, 2, 4, 6, 8, or 10 cents. The gamble gives you a 50% chance of doubling what you bet and, since you are uninsured, a 50% chance of losing half of what you bet.

If you had insurance, you would not have lost anything in case you gambled and lost.

They then received these instructions and could only proceed if they indicated they understood the gamble:

Here are two examples of how the bet may work out.

Example 1: if you choose to bet 6 and you <u>win</u>, you will receive:

***10 cents** for completing the HIT*

***4 cents** that you did not use to bet*

*6 cents that you bet x 2 = **12 cents** as your bonus*

total = 26 cents

Example 2: if you choose to bet 6 and you <u>lose</u>, you will receive:

***10 cents** for completing the HIT*

***4 cents** that you did not use to bet*

You are uninsured so:

*6 cents that you bet divided by 2 = **3 cents** as your bonus*

total = 17 cents

Remember that if you had had insurance you would have not lost anything of what you bet (last sentence in red).

On the next page, participants in both conditions indicated how much they wanted to bet.

After placing their bet, participants in all conditions learned their outcomes and were told that they would be paid accordingly within 7 days. To introduce the final attention check they then saw the gamble again and read the following text:

You are almost done. We are just checking to make sure you actually understood the gamble. Please note that your answers on the questions below will NOT affect your payment.

[picture of gamble]

Remember that you started out with a 10 cent bonus.

Imagine that you bet 4 cents on the gamble and won.

All participants then answered the following three questions (1) “How many cents did you keep (not bet)?” (2) “What would be the payoff for betting 4 and winning?” (3) “How much would your total earnings be in this case (including the 10 cent payment for completing the hit)”. Finally, they indicated their gender, age, WorkerID, and how they found the hit. They could also leave a comment if they wished.

Chapter 4 - Incomplete understanding of insurance facilitates fraud acceptance

Abstract. This paper documents how people think about insurance and proposes that most policyholders feel insurance is a bad investment. When asked to describe insurance, many people mention getting (direct or indirect) payments from the insurance company. When there is no return-on-investment on premiums paid (i.e., no approved claims), people feel that the money spent on insurance is wasted. We propose that this leads policyholders to believe that filing illegitimate claims is a way of compensating for the perceived imbalance between premiums paid and returns obtained. In sum, we suggest that conceptualizing insurance as getting payments from the insurance company contributes to positive attitudes towards insurance fraud.

This chapter is based on: Van Wolferen, J., Inbar, Y., & Zeelenberg, M. (2014) Incomplete understanding of insurance facilitates fraud acceptance.

Insurance fraud is a big problem, but unfortunately nobody knows quite how big. By definition, successful fraud goes unnoticed, so estimates of its occurrence are always imprecise. Still, the available evidence clearly indicates that fraud is a non-negligible problem. For example, the Coalition Against Insurance Fraud (CAIF; 2013) states that insurance fraud 'steals at least $80 billion every year' in the U.S. alone, although they do not specify what percentage of the total premium volume this constitutes. Reports from specific insurance domains paint a similarly grim picture. For example, the European Healthcare Fraud & Corruption Network estimates the annual cost of health care fraud to be €56 billion in Europe and €180 billion globally (Gee, Button, Brooks, & Vincke, 2010). According to the same report, European countries spend 1 trillion on health care every year. If this fraud estimate were correct, 5.6% of total claims in health care would be fraudulent. Similarly, a report based on data from the British Association of Insurers finds that detected fraud in household insurance amounted to £983 million in 2011, which accounts for 5.7% of the total claim volume (Button, Pakes, & Blackbourn, 2013). In addition, the report suggests that £2 billion worth of fraudulent claims might go unnoticed in the household insurance sector annually.

A different way of obtaining estimates of the seriousness of insurance fraud is to ask insurance-company employees what percentage of claims they believe to be fraudulent. A study that surveyed employees at U.S. insurance companies (Insurance Services Office, 2000) found that 42% of respondents thought that more than one-fifth of total claim volume was due to "soft" fraud—that is, increasing the amount requested on otherwise legitimate claims. In addition, 6% of respondents thought that more than one-fifth of total claim volume was due to "hard" fraud—that is, making up or staging accidents (definitions of hard and soft fraud by the Insurance Information Institute, 2010). In a different survey of 143 U.S. insurers, 45% of respondents thought 5-10% of the total claim volume was fraudulent. Almost a third of respondents (32%) indicated that fraud might even be 20% of total claim volume (Insurance Journal, 2012). These survey estimates—even though they differ substantially from the numbers in the previous paragraph—show that insurers agree that fraud costs a considerable amount of money. Similarly, the inception of organizations like the National Insurance Crime Bureau illustrate that fraud is big enough of a problem for insurers to unite and fight it.

It is also noteworthy that insurance fraud is deemed such an acceptable crime. In a U.K. sample of the general public, 37% percent of respondents said they 'would not rule out committing insurance fraud in the future' and 29% said that would be 'acceptable or borderline'. Worse, 47% would not rule out 'exaggerating an insurance claim' and 40%

would deem that 'acceptable or borderline' (Button et al., 2013). Similarly, the Coalition Against Insurance Fraud concludes that 'a growing number of Americans view insurance fraud as a minor crime and no big deal' (2007, p. 14). Why might insurance fraud be seen as so acceptable?

The problem of insurance fraud has generally been approached from an expectancy value perspective (Becker, 1968) that focuses on audit probabilities, fines, the costs of verifying and falsifying claims, and the technical aspects and deterrent effects of detection (e.g., see the special issue in the Journal of Risk and Insurance introduced by Derrig, 2002). Similarly, models used to determine the optimal auditing strategy assume risk averse expected utility maximizing policyholders (Dionne, Giuliano, & Picard, 2008) and some investigations gauge whether insurers' auditing practices confirm to theoretical predictions (e.g., Tennyson & Salsas-Forn, 2002). These and other findings have contributed greatly to the fight against insurance fraud by providing insurance fraud prevention units the tools they need to detect and deter fraud.

However, when the focus is on the detection and deterrence of fraud through audits and mathematically sophisticated analyses of claims, one question is usually not addressed: *Why* do people think insurance fraud is acceptable? In most, if not all, models of insurance fraud, the decision to commit insurance fraud is assumed to be a function of the probability of success times the expected costs and benefits of filing a fraudulent claim. Some models do incorporate "morale costs" of committing insurance fraud (e.g., Picard, 2000) but it is unclear what determines their size. There have only been a handful of articles that attempt to identify factors that determine attitudes towards insurance fraud; we review them below.[22]

Why do people think insurance fraud is acceptable?

To figure out why people think insurance fraud is acceptable, it is interesting to look at the excuses people use to people legitimize it. Additionally, these excuses turn out to give insight into what people think the return on insurance payments should be. In the literature, as well as in casual conversations with friends and colleagues, one frequently-mentioned justification indicates that people feel like they are not getting their money's worth when they buy insurance: many people say that filing fraudulent claims can be a way of "seeking a fair return" on paid premiums (e.g., Button et al., 2013; Coalition Against Insurance Fraud, 2007; Verschuur, 1992). Other studies find that claim padding (increasing the amount of an otherwise legitimate claim) can be justified as "making up for a deductible or past premiums"

[22] We have summarized some of those findings in another paper that reviews recent empirical evidence for moral hazard in the insurance industry. Interested readers are referred to van Wolferen, Inbar, and Zeelenberg (2013).

(Tennyson, 1997, 2002). In fact, in one study people thought it was more acceptable for someone to commit insurance fraud when they had a high, compared to a low, deductible (Miyazaki, 2008).

Justifying insurance fraud as making up for past or current costs and premiums highlights the perceived imbalance between the insured's input and the insurer's output. It is illustrative of the perception that insurance companies take a lot of money, but provide nothing in return. Generally, when people buy a product or service, or invest in something, they expect some sort of return (e.g., Fiske, 1991) and people may feel like they are not getting enough in exchange for their insurance payments. Put differently, policyholders may see fraud as a way to get the return-on-investment they deserve for paying insurance premiums. Clearly, people can only think insurance fraud is a way of balancing the scales when they feel that they are not getting enough value for money. Why do people perceive their relationships with insurers to be so inequitable?

We propose that one reason is that policyholders think about insurance in a way that is different from how students of insurance decisions and policymakers think about insurance. Historically, insurance has been a way for groups of people to reduce risk by sharing it (Bernstein, 1998). Pooling and redistributing money to those in need protects everyone in the insurance agreement against financial disaster.[23] Today, when thinking about insurance, we suggest that most people do not feel like they reduce their financial risk by pooling their money with that of others. Instead, they may focus on how much the insurance company will provide in return for their premium payments. Consequently, when the perceived return on investment is small or absent, people feel like they have not received their money's worth—that they have wasted money on insurance.

We are not the first to suggest that policyholders generally misunderstand insurance (although as far as we know, we are the first to examine the perception of insurance empirically). Clarke (1990, p. 3) notes that "the insured's familiarity with the purposes, values, and conditions of insurance cannot any longer be taken for granted by insurers" and that insurance companies are likely to be perceived as "fair game." Viaene and Dedene (2004) note that insurers should be careful not to assume that policyholders' knowledge of the workings and goals of insurance is correct. Finally, in a recent book Kunreuther, Pauly, and McMorrow (2013) thoroughly describe how a misunderstanding of insurance may contribute

[23] Most consumer insurance policies like travel, health, and home-owner insurance are based on risk-pooling schemes. Other types of insurance policies exist where risk is transferred from one person to one other wealthy person or entity.

to different demand-side anomalies in insurance markets. Throughout the book, the authors repeat that most people do not realize that the essence of insurance is risk reduction through risk pooling. A failure to recognize this core feature of insurance is argued to lead to frustration and sometimes even to policy cancellation. Even though insurance reduces the risk that the policyholders face, policies are often cancelled after some time because they are perceived to be a bad deal—an investment that does not pay off (for an example in flood insurance, see Michel-Kerjan, Lemoyne de Forges, & Kunreuther, 2012). Together, these findings thus suggest that people think of insurance as an investment. And, as with all investments, insurance is only perceived to be a good investment if it pays off.

We start our investigation of how ordinary people think about insurance and insurance fraud with determining whether they understand the core elements of insurance, namely that insurance is a risk pooling mechanism. In addition, we aim to advance a positive account of how people think about insurance.

Incomplete understanding of insurance and attitude towards fraud

The main contribution of this paper is the finding that an incomplete understanding of insurance leads to greater acceptance of insurance fraud. We propose that fraud acceptance is facilitated when policyholders think about how much money the insurance company returns in exchange for paid premiums. A focus on these 'returns' leads not receiving payments from the insurance company to be perceived as a loss. Furthermore, policyholders feel that they are spending their money in vain—that is, that they are wasting money on something that only rarely provides a benefit. If that benefit is the most salient or important feature of insurance, its absence will lead policyholders to think getting insurance was a bad investment. We propose that one way to resolve this is to claim money that they are not legally entitled to (for a similar model of tax evasion, see Wenzel, 2002).

Based on this reasoning, we expected that the more strongly people feel that insurance does not provide enough of a return on investment (that money spent on insurance is wasted), the more they will find insurance fraud—which creates such a return—acceptable. We test this proposition in Studies 1a and 1b. We also expected that certain situational factors would increase the perception of insurance premiums as money wasted. One obvious such factor is the salience of the premium payments. To test this idea, in Studies 2a-3 we remind people of the price of insurance and find that this increases their acceptance of insurance fraud. In the method and results sections below we report the most important details of our studies. In the Appendix we report additional measures and analyses. In sum, we report how we determined our sample sizes, all data exclusions (if any), all manipulations, and all measures in the study.

Pilot studies: What do people think insurance is?

Here, we report data obtained on four different occasions. We first asked 40 Tilburg University (TiU) students and 40 people in a shopping mall in Tilburg, the Netherlands to explain in their own words what they thought insurance was and how it works. Subsequently, two samples of U.S. participants recruited via Amazon's Mechanical Turk, an online labor marketplace (MTURK; Buhrmester, Kwang, & Gosling, 2011), answered the same question in return for a small cash payment. Demographics and key results are shown in Table 1.

Table 1. Demographic information and coding of answers to the question: 'what is insurance?' from four samples in the Pilot Studies.

Sample (*N*)	M_{age}, [range], % female.	"getting money back" (count / %)	"risk sharing" (count / %)
TiU students (*N*=40)	19.28, [18-26] 90.0%	36 (90.0%)	1 (2.5%)
Tilburg mall (*N*=40)	52.33, [18-77] 45.0%	19 (47.5%)	5 (12.5%)
U.S. MTURK (*N*=104)	30.64, [18-67] 42.3%	76 (73.1%)	12 (11.5%)
Study 1a (*N*=157)	35.08, [18-67] 51.6%	127 (80.9%)	22 (14.0%)
Total (*N* = 341)		258 (75.7%)	40 (11.7%)

Materials and procedure

The Tilburg samples (students and mall-goers) were given a questionnaire asking them to "describe in [your] own words, what health insurance is and how health insurance works". They also answered the questions "How often do you deal with insurance in your daily life?" (1= almost never, 7 = on a daily basis), "How much do you know about insurance?" (1 = next to nothing, 7 = a lot), "What is your perception of insurance and insurance companies?" (1 = very negative, 7 = very positive). [24]

The two U.S. samples were asked, "Please describe in your own words what insurance is and how insurance works. If you find it hard to describe insurance, please write down your answer as if you were describing insurance to a 16 year-old." Many people referred to health and or auto insurance to explain what insurance is.

Results

The first author coded the data from the two Tilburg samples. Initially, the goal was to determine how many people mentioned the principle of risk sharing in their description of insurance. After having read all descriptions, it became clear that many people focused on the insurer making payments to the insured. Explanations of when and how the insurer would make a payment were very frequent. In those descriptions, the implicit assumption is that the expected 'return-on-investment' comes in the form of direct or indirect payments. To determine how many of the respondents mention getting money back from the insurer, the first author coded the data from the Tilburg samples again.[25]

To obtain a more objective measure of how respondents described insurance, two coders coded the 261 texts obtained in the two other samples.[26] The coders used the following rule to determine whether the insurance description mentioned "getting money back" (1) or not (0): "Code '1' if a payment from the insurer to or for the insured is mentioned. Examples: "it'll cover expenses" or "they pay for your damages", otherwise code '0'." To determine whether insurance was explained as 'sharing risk' (1) or not (0), the coders used the following rule: "Code '1' if insurance is described as spreading/sharing risk. Examples: "it's a way for people to share the risks they face" or "insurance is a group of people pooling money together", otherwise code '0'." The coders initially agreed in 81.5% of the cases when coding

[24] These questions were given to participants in Dutch. We thank Job Krijnen for setting up this study and helping with the data collection in Tilburg.

[25] The data from relatively few subjects in the mall sample fit that description. We think this is because a significant proportion of that sample wrote about what is wrong with the current system and how it should be improved, instead of answering the question.

[26] We thank Shere Nahari Santomé and Sebastiaan Verhulst for coding these texts.

the ‘money back’ variable and in 96.5% of the cases when coding the ‘risk sharing’ variable. Disagreements were resolved through discussion with the first author.[27]

The results of these pilot studies suggest the majority of respondents focused on what the return would be on the money they paid the insurance company. Importantly, the return was described as a payment from the insurance company to the policyholder. The following is an example of how approximately seventy-five percent of the sample described insurance as a way to cover future expense:

> *Insurance basically is some sort of way to make sure that if something happens to you, you have a way of paying for that. Like, if you have to go to the doctor, you can pay for it with insurance. You make payments for insurance and it makes sure to keep you covered...*

One participant put it more succinctly:

> *I pay them and if I have a problem, they will eventually pay me.*

These results also indicate that many people do not spontaneously think about the collective nature of insurance. That is, only a few respondents describe insurance as a way to reduce risk by sharing it with other policyholders. The following is an example of how approximately ten percent of the sample described insurance as risk-sharing:

> *Insurance is a way to make sure that if something unexpected happens, you will be able to afford it. Every month you pay money to an insurance company, which pools it together with everyone else's money. Then, if you are hit with an expense covered by the insurance (say, a car accident or an illness), the insurance company will use money from that pool to pay for you.*

If policyholders indeed mainly think about the possible monetary return on their insurance payments, it is inevitable that some might feel that the return-on-investment is not large enough. Consider the case when a person is lucky enough to never experience misfortune that needs covering by the insurance company. Arguably, this is a best-case

[27] Typically, it is assumed that people buy insurance to reduce the financial risk they face. It was hard to determine whether or not people described insurance as reducing financial risk. To determine whether insurance was explained as risk-reduction (1) or not (0), the coders used the following rule: “Code ‘1’ if insurance is described as reducing financial risk. Examples: 'insurance takes away the possibility that you have to pay a large sum of money', and 'it reduces the financial risks that you face', otherwise code ‘0’. When in doubt, only code ‘1’ if the insured and uninsured situation are explicitly compared financially. The coders agreed in 81.2% of the cases and disagreements were resolved through discussion with the first author. This narrow formulation of risk-reduction fit 68 (26.1%) of the answers. However, if we employ a more flexible formulation and also code answers like ‘insurance guarantees you will be compensated if something happens’, ‘it saves you money if there is actual damage’, ‘it protects you when you get into an accident’, or ‘insurance prevents bankruptcy’, 187 (71.6%) cases fit the description.

scenario. However, policyholders who pay monthly premiums but never receive ‘returns’ might ironically feel like they got the bad end of the deal. Not receiving any payments from the insurance company might lead money spent on insurance premiums to be seen as ‘wasted’. This perception of insurance does not preclude that policyholders who *do* receive a return can also feel taking out insurance was a bad idea. For policyholders who construe insurance as an investment that should provide something in return, the perceived return on investment might still be too low in case the insurance company sends (too) few payments.

We propose that the perceived lack of return-on-investment facilitates tolerant attitudes towards insurance fraud. Specifically, we think that the smaller the perceived return on investment on insurance payments, the more likely people are to think insurance fraud is acceptable. To test this proposition, we measured the extent to which people felt that their insurance money was wasted if they did not require insurance reimbursements. We correlated that with how acceptable they thought insurance fraud was. We find that the more people think money spent on insurance without filing a claim is money down the drain, the more they think fraud is acceptable.

Study 1a: Wasted insurance payments and attitude towards insurance fraud.

Method

In this study we formulated four questions designed to measure to what extent people think their insurance payments are wasted if they do not file insurance claims. We think this feeling is associated with a greater perceived acceptability of insurance fraud.

Participants. One hundred and fifty seven people completed the study on MTURK (M_{age}= 35.09, range 18-67; 91 female) in exchange for \$0.50 on September 26, 2012. People could only participate if they lived in the U.S.

Materials and procedure. The study was programmed in Qualtrics and we included an instructional manipulation check (IMC) to weed out inattentive participants (see Oppenheimer, Meyvis, & Davidenko, 2009). Participants were not allowed to start the study if they failed the IMC.

Participants completed three parts of the survey in random order. One part asked them to describe what they thought insurance was. These data are reported in the Pilot Studies. In a second part, they were asked to what extent they thought their insurance money was wasted if they did not file claims. Specifically, they rated their agreement with the statements “If I paid insurance premiums for 10 years but never filed a claim, I would feel annoyed to have wasted my money”, “I think my insurance premium payments will have been wasted if it turns out I never have losses that need covering”, “The money I spend on insurance is money down the drain if it turns out I never require the coverage", and "If I paid monthly insurance premiums for an extended period of time, but never have to file a claim, I would feel like getting insurance was a bad deal" on a 7-point scale (1 = Strongly disagree, 2 = Disagree, 3 = Somewhat disagree, 4 = Neither agree nor disagree, 5 = Somewhat agree, 6 = Agree, 7 = Strongly agree), $\alpha = .97$. We use the term 'waste-scale' to refer to these items in the remainder of the paper.

Another part of the survey measured participants' attitude towards insurance fraud. Participants rated the extent to which they found each of the following behaviors acceptable: "Misrepresenting the facts on an insurance application in order to obtain insurance", "Misrepresenting the facts on an insurance application in order to obtain a lower rate", "Submitting an insurance claim for damages that occurred prior to the accident being covered", "Inflating an insurance claim to help cover the deductible", "Misrepresenting the nature of an accident to obtain insurance payment for a loss not covered by the policy", "Falsifying receipts to increase the amount of the insurance settlement", and "Deliberately overestimating losses to increase the amount of the insurance settlement". Each of these items was rated on a 7-point scale (1 = Very unacceptable, 2 = Unacceptable, 3 = Somewhat acceptable, 4 = Neutral, 5 = Somewhat acceptable, 6 = Acceptable, 7 = Very acceptable) and the scale had good reliability, $\alpha = .95$. Some of these questions were taken from Tennyson (2002). We use the term 'fraud-scale' to refer to these items in the remainder of the paper.

We also included a measure of how entitled people feel to health care. At the time we were interested in this variable and presented it to participants after they completed the three elements above. For completeness, we report those data here too. Participants rated the extent to which they agreed with each of the following items on the same scale as the 'waste-questions': "I think my insurance policy should cover the best treatments available, regardless of the cost", "If I am seriously ill, the doctor should try everything to save my life no matter how much it might cost", "I think the doctor gets paid enough to spend more time with me than he currently does", "Even if there is just a 1% chance that a really expensive treatment works, my insurance company should allow me to take it", "I think my insurance policy should cover whatever I and my doctor decide to do, even if that is extremely expensive", "My insurance company should give me access to new treatments–even really expensive ones–as soon as they are proven effective", and "I should be able to receive new or experimental treatments even if they are very expensive and it is not certain that they are effective". Each of these items was rated on a 7-point scale (1 = Strongly disagree, 2 =

Disagree, 3 = Somewhat disagree, 4 = Neither agree nor disagree, 5 = Somewhat agree, 6 = Agree, 7 = Strongly agree) and the scale had good reliability, α = .90.

Finally, participants indicated their gender, age, MTURK worker ID, and their political stance using the following question: “When it comes to politics, do you usually think of yourself as liberal, moderate, conservative, or something else?” (1 = Very liberal, 2 = Liberal, 3 = Slightly liberal, 4 = Moderate/middle-of-the-road, 5 = Slightly conservative, 6 = Conservative, 7 = Very conservative, 8 = Libertarian, 9 = Don’t know / Not political, 10 = Other, with possibility to enter text).[28]

Results

In Table 2, we report the correlations between the main variables in the Study. The most important result is that we find a positive correlation between the extent to which people think their money is wasted if unused, and how acceptable they think insurance fraud is (r = .282, p <.001).

Table 2: Means, standard deviations, and (non-)parametric correlations between variables in Study1a (N = 157).

	M (SD)	Waste	Fraud	Entitlement
Waste	3.97 (1.91)	α = .97	r =.282, p < .001	r = .019, p = .812
Fraud	2.38 (1.26)	ρ =.291, p < .001	α = .95	r = -.115, p = .151
Entitlement	5.30 (1.11)	ρ =.012, p = .884	ρ = -.134, p = .095	α = .90

Note: The diagonal displays Cronbach’s alpha, r = pearson correlation coefficient, ρ = spearhman-rho correlation coefficient. The 7-item fraud scale ranged from 1 = Very unacceptable to 7 = Very acceptable. The 4-item waste scale and the 7-item entitlement scale ranged from 1 = Strongly disagree to 7 = Strongly Agree. Higher numbers indicate greater tolerance of fraud, stronger feelings that money spent on insurance is wasted if ‘unused’, and greater sense of entitlement.

[28] We find no significant correlations between political orientation (1-7), and fraud tolerance and entitlement.

The correlation is a first indication that the feeling of spending money on insurance without a return on investment in claims paid is associated with more positive attitudes towards insurance fraud. However, we cannot make inferences about the causal direction of this effect. In addition, we might have overestimated the correlation between the waste- and fraud-scale because all items in both scales were positively framed (e.g., higher score is more agreement and people who are 'agreeable' would spuriously boost the size of the correlation; Heiman, 2002).[29] Therefore, in an attempt to replicate the result in Study 1a, we negatively frame two questions of the waste scale and test if we find the same correlation. In addition, we tried to get more insight into which features of insurance contribute to the perception that money spent on insurance is money down the drain. This helps establish possible manipulations that we could use in follow-up studies in which we test whether the effect of the feeling of having wasted money on attitudes towards fraud is causal.[30]

Study 1b

Participants. One hundred and seventy eight people completed the study on MTURK (M_{age}= 33.20, range 18-67; 60 female) in exchange for $0.35 on April 5, 2013. People could only participate if they lived in the U.S.

Materials and procedure. The study was programmed in Qualtrics and included an IMC. Participants were not allowed to start the study if they failed the IMC.

Participants completed three parts of the survey in random order. One part presented the fraud-scale (α = .93) and another part presented the waste-scale that we used in Study 2a. However, the framing of two items was reversed such that they read "I think my insurance premium payments are *well spent* if it turns out I never have losses that need covering" and "If I paid monthly insurance premiums for an extended period of time, but never have to file a claim, I would feel like getting insurance was a *good deal*." These items were reverse-coded in the analyses and the scale was reliable (α = .84).

A third part of the survey required participants to indicate on a 7-point scale to what extent six features of insurance made them feel like insurance was a bad (1) or a good (7) investment. These features were "having to pay a deductible", "having a claim approved by my insurance company", "having a claim denied by my insurance company", "having to pay

[29] We thank Linda Babcock for pointing this out.
[30] Note that we ran study 1b after we ran Study 2a, 2b, and 3. It turns out that in these three studies, we intuitively picked a good manipulation ('having to pay (monthly) premiums') that was associated with feelings of wasting money on insurance.

monthly premiums", "knowing I am covered in case of an accident", and "having to make a co-payment for a visit to the doctor, medication, or medical procedure".

Finally, participants indicated their gender, age, WorkerID, through which website they had found the survey, and were allowed to leave a comment if they wished.

Results

We replicate the result of Study 1a and find that answers on the waste- and fraud scale are correlated, r = .230, p = .002; ρ = .209, p = .005. The second goal of this study was to establish which features of insurance make buying insurance seem like a bad investment. Means and standard deviations are given in Table 3, along with how the variation in the answers on these items is associated with answers on the fraud- and waste-scale. The means of items 1, 4, and 6 are higher than we expected but we find that an overwhelming majority uses a '4' (midpoint of the scale) or lower to indicate how these features affect whether getting insurance is a bad (1) or a good (7) investment. The most important result is that 83% of participants felt that paying monthly premiums made insurance feel like a bad investment, and that the more they felt this way, the more tolerant they were of insurance fraud (r = -.21, p = .004).

Table 3. Means, standard deviations, the percentage of people answering 4 or lower, and correlations with other measures for bad/good investment rating for each of six features of insurance in Study 1b, $N = 178$ for all items.

	M (SD)	≤ 4	Correlation with fraud-scale [a]	Correlation with waste-scale
1. Having to pay a deductible	3.15 (1.34)	88.2%	$r = -.126$, $p = .093$	$r = -.341$, $p < .001$
2. Having a claim approved by my insurance company	5.89 (1.24)	14.6%	$r = .022$, $p = .771$	$r = -.113$, $p = .133$
3. Having a claim denied by my insurance company	1.61 (0.84)	99.4%	$r = .017$, $p = .821$	$r = -.208$, $p = .005$
4. Having to pay monthly premiums	3.37 (1.15)	89.3%	$r = -.212$, $p = .004$	$r = -.375$, $p < .001$
5. Knowing I am covered in case of an accident	6.20 (0.88)	5.6%	$r = -.148$, $p = .049$	$r = -.217$, $p = .004$
6. Having to make a co-payment for a visit to the doctor, medication, or medical procedure	3.66 (1.40)	72.5%	$r = -.128$, $p = .088$	$r = -.361$, $p < .001$

[a] One way to interpret the negative correlations is as follows: the more a feature of insurance makes someone feel like insurance was a bad investment, the more acceptant someone is towards fraud.

In the studies that follow we test whether there is a causal link between feelings of wasting money on insurance and attitudes towards fraud. In Study 1b, we have established that having to pay premiums feels like wasting money on insurance for many people. Therefore, in the following studies, we use the price of insurance as a between-subjects manipulation. More specifically, we give half our participants information about how much insurance costs but keep this information from the other half. The idea in these studies is that if having to pay premiums makes insurance feel like a bad investment, making premium payments salient may lead people to find the idea of recouping their losses by filing fraudulent claims more acceptable. We thus predict that subjects who are reminded of the price of insurance will find fraud more acceptable than subjects who are not reminded of the price of insurance. In the next section, we describe two experiments (Studies 2a and 2b) that have highly similar methods. Therefore, we describe the method extensively in Study 2a and only indicate the differences in the method sections of Study 2b.

Study 2a: Travel insurance premium (€65) and attitude towards fraud

Method

Participants. Fifty-two students (M_{age}= 18.94, range 17-24; 23 female) at the Fontys University of Applied Sciences in Tilburg participated in a 20-minute session of unrelated experiments in exchange for 5 euros in October 2011. They were randomly assigned to one of two (premium vs. control) conditions.

Materials and procedure. The study was programmed and administered using Authorware 7.0 and demographic data were collected at the beginning of the session. One screen indicated the start of every study and upon starting the current study all participants were asked to imagine the following situation:

> *You have just returned from your travels to a sunny country. You lost your digital camera during this holiday. You are not sure whether you accidentally left your camera somewhere or it was stolen. You have travel insurance so you could report the camera as stolen. This would allow you to buy a new camera with the insurer's money.*

Following that, the participants in the premium condition read:

> *You pay 65 euros per year for the insurance.*

In the control condition, there was no additional text. Next, all participants answered one question: 'In this case, how acceptable is it to report the camera as stolen?' on a 7-point scale (1 = Very unacceptable, 7 = Very acceptable).

Results

A *t*-test ($t(50) = 1.86$, $p = .069$, $d = 0.53$) and a Mann-Whitney *U*-test ($U = 425$, $z = 1.69$, $p = .092$, $r = .23$) indicate that participants in the premium condition ($n = 28$; $M = 5.39$, $SD = 1.29$; *Mean Rank* = 29.68) thought it was more acceptable to report the camera as stolen than participants in the control condition ($n = 24$; $M = 4.71$, $SD = 1.37$; *Mean Rank* = 22.79). However, possibly because of the low *N*, both these tests are only marginally significant so we sought to replicate the study with more power. In addition, we also wanted to examine the effects in an adult population instead of a student population. To this end, we ran Study 2b on MTURK.[31]

[31] We also ran one study with 101 participants (M_{age}= 33.41, range 18-67; 53 female) on MTURK where we just changed €65 into $65 and did not find an effect ($t(87) = 0.70$, $p = .485$, $d = .15$; $U = 1083.50$, $z = 0.82$, $p = .412$, $r = .08$). In retrospect, we found out $65 is a very good deal for annual travel insurance so we used a more realistic price of insurance in Study 2b.

Before running study 2b, we looked up several quotes via www.squaremouth.com, a site designed to compare travel insurance providers and policies. The price of the top 15 offers for annual travel insurance for 1 person who would have 2 trips per year ranged between \$44 and \$267 and included 11 policies that cost more than \$150 (in October 2012). Therefore, we replicated the design of Study 2a but told participants in the premium condition that they paid \$150 annually.

Study 2b: Travel insurance premium (\$150) and attitude towards fraud

Method

Participants. Three hundred participants (M_{age}= 28.49, range 18-71; 98 female) completed this study on MTURK in exchange for \$0.10 in October 2012. People could only participate if they lived in the U.S. They were randomly assigned to one of two (premium vs. control) conditions.

Materials and procedure. The study was programmed in Qualtrics and an IMC at the end of the survey (instead of at the beginning, as had been the case in previous studies). This allowed us to test whether the effect of our manipulation was different for people who failed versus passed the IMC (it was not; see Appendix). The scenario, manipulation, and DV were the same as in Study 2a. Because we used a variant of this scenario in other MTURK studies (see footnote 31) we asked whether people recognized the scenario and excluded them from the analyses if they did.

Results

Twenty-two people indicated they had seen the question in other surveys and four people took less than 10 seconds to read the scenario and complete the DV. We exclude their answers from the analyses and are left with data from 275 participants.[32]

A t-test ($t(273) = 2.95$, $p = .003$, $d = 0.36$) and a Mann-Whitney U-test (U = 11456.50, $z = 3.13$, $p = .002$, $r = .19$) indicate that participants in the premium condition (n = 142; $M = 5.02$, $SD = 1.52$; *Mean Rank* = 152.18) thought it was more acceptable to report the camera as stolen than participants in the control condition ($n = 133$; $M = 4.48$, $SD = 1.51$; *Mean Rank* = 122.86).

In Studies 2a and 2b, we find that reminding people of the price of insurance leads them to think reporting a camera as stolen—while it might be lost or stolen—more acceptable compared to when they are not reminded of the price of insurance. To be fair, some travel insurance policies actually cover 'lost items' and one might argue that we did not really

[32] Including their answers in the analyses does not meaningfully change the results.

measure how people thought about filing illegitimate claims. Therefore, in Study 3, we replicate the design of studies 2a and 2b but ask participants to evaluate an unequivocally fraudulent claim.

Study 3: Health insurance premium ($3,500) and attitude towards fraud

Method

In preparing Study 3, we again wanted to present people with a reasonable price of insurance. We looked up the average cost of health insurance in the U.S. and found that premiums have been rising dramatically in the past 10 years (Kaiser Family Foundation, 2012). The average annual premium was $4,316 in 2011, a 97% increase compared to 2002. Conversations with American colleagues suggested that $4,316 might seem like a high price to many participants. Therefore, we presented a health insurance premium that was a bit lower, namely $3,500. This preempts the argument that we would find an effect only because the people in our experiment felt like we presented them with an extraordinarily bad deal. Consistent with Studies 2a and 2b, we expected that reminding people of the price of insurance would lead them to think insurance fraud is more acceptable.

Participants. Three hundred and one participants (M_{age}= 27.91, range 18-64; 130 female) completed this study on MTURK in exchange for $0.10 in October 2012. People could only participate if they lived in the U.S. They were randomly assigned to one of two (premium vs. control) conditions.

Materials and procedure. The study was programmed in Qualtrics and we included an IMC at the end of the survey, again, performance on the IMC did not change the effects in this study (see Appendix). All participants were asked to imagine the following situation:

Your health insurance policy does not cover the treatment that your physician needs to give you. It does cover a slightly different treatment and the physician—in his report to the insurance company—could pretend that he gave you the covered treatment.

Following that, the participants in the premium condition read:

You pay $3500 annually in health insurance premiums.

In the control condition, there was no additional text. Next, all participants answered one question: ‘In this case, how acceptable do you think it is to have the physician report the covered treatment instead of the uncovered treatment?’ on a 7-point scale (1 = Very unacceptable, 7 = Very acceptable). At the end, all participants answered the question whether they had seen the physician question in other surveys (yes/no), whether they thought $3500 was a low/normal/high price for health insurance, followed by the IMC.

Results

Forty-one people indicated they had seen the question in other surveys and 33 people took less than 10 seconds to read the scenario and complete the DV. We exclude their answers from the analyses and are left with data from 237 participants. In the supplemental materials we provide the tests of how these outlier criteria affect the main results. Overall, the effect of presenting the premium on attitudes towards fraud is robust.

A t-test ($t(235) = 2.59$, $p = .010$, $d = 0.34$) and a Mann-Whitney U-test ($U = 8405.50$, $z = 2.68$, $p = .007$, $r = .17$) indicate that participants in the premium condition ($n = 117$; $M = 5.08$, $SD = 1.77$; *Mean Rank* = 130.84) thought it was more acceptable to file a fraudulent claim than participants in the control condition ($n = 120$; $M = 4.45$, $SD = 1.93$; *Mean Rank* = 107.45). These findings indicate that a reminder of the price of insurance leads people to be more tolerant of an act that is clearly illegal. In the following section, we summarize and elaborate on our findings.

General discussion

The understanding of insurance by members of the general public is different than that of insurers and insurance researchers. We believe that the way in which the general public understands and evaluates insurance provides valuable insights into perceptions of insurance fraud. The studies presented here found that few people think of insurance as a way of reducing risks by sharing them with others. Instead, when asked to describe insurance, most people focus on the possibility of getting some money back from the insurer. The research also revealed that the more people feel like they are wasting money on insurance if they never get anything back, the more tolerant they are towards insurance fraud. The final three studies found that people also think insurance fraud is more acceptable when they are reminded of the (high) price of insurance. These findings have implications for the literature on insurance fraud, as well as for insurers who might be interested in reducing insurance fraud. We review both areas below.

Insurance as a bad investment?

It has been proposed that people treat insurance as if it were an investment (e.g., Michel-Kerjan et al., 2012). We suggest that this perception is especially likely in retrospect, and that the most common perception is that insurance is a *bad* investment. Perceiving insurance as an investment *in prospect* (i.e., at the time of signing a contract), implies an expectation that taking insurance will be profitable; that means that one will make many small periodic payments in the hopes of receiving a future return that is large enough to yield a net gain. To us, it seems unlikely that people think this way—and indeed only one of the participants in our pilot studies described insurance as a gamble in which one bets on making a net profit.

In contrast, treating insurance as an investment in *retrospect* implies that some time after having signed an insurance contract, people compare what they have paid so far with what they have received. Focusing on the monetary return, rather than the return in the form of risk reduction, would in most cases lead to the conclusion that insurance was a bad investment. Imagine best-case scenarios where people have not experienced sickness, accidents, or any other insured mishap. The insured will have made payments to the insurer for some time, but the insurer will have returned exactly $0. In this situation, it is very tempting to conclude that insurance was a bad investment indeed.

Insurers thus sell products that nobody hopes to use—but when this best-case scenario happens, clients are unhappy. In the remainder of this text, we elaborate on possible solutions for this problem. One prediction that follows from the studies in this paper is that attitudes towards insurance fraud become less tolerant as policyholders feel they receive an appropriate return for their payments. We think there are at least two ways in which insurers could proceed. First, policyholders could be educated about the collective nature of insurance. Specifically, as Kunreuther et al. (2013, p. 231) propose: "... if [policyholders] do not suffer a loss, they should celebrate their good fortune in being insured rather than perceive their premiums as wasted." Given the abstract nature of risk reduction through risk sharing, it might be hard to convince all policyholders that they really are getting something in return for the money they pay insurers. In addition, advertisements that claim clients should be happy if they never receive a payment from the insurance company might prove counterproductive. A second solution would focus on fixing the symptoms (that is, high tolerance of insurance fraud) while leaving the incomplete understanding of insurance intact. Specifically, insurance companies could implement no-claim bonuses. That is, at the end of the year, insurance companies return some of the 'unused' money to clients who have not filed a claim. This would reduce these clients' feeling that the money they spent on insurance is wasted, which in turn could reduce tolerance of insurance fraud. The problem associated with this solution is that insurers would have to increase premiums to be able to afford such a scheme, and that it does not solve the problem at its root.

A third solution again builds on the idea that people feel like they are not getting enough in return for the premiums they pay. This solution requires insurers to educate policyholders about what it is exactly that they pay premiums for. Essentially, people pay a small amount of money to reduce their (financial) risk. In lay terms, insurance means buying ‘peace of mind’. If insurers could continuously remind or convince their policyholders that the return on investment comes in the form of peace of mind, the perception of money down the drain might be reduced, which in turn would reduce tolerance towards insurance fraud. Risk-reduction and peace of mind are abstract constructs that are not easily quantifiable, so this solution is not easily implemented. However, simple ads that state buying insurance means buying peace of mind might be a first step in the right direction.

Finally, although we have identified one cause of tolerant attitudes towards insurance fraud, it is unclear to what extent these findings apply to actual insurance fraud. Surely, other factors like the perceived likelihood of getting away with fraud and the consequences of getting caught affect this decision. Nevertheless, by understanding why people think insurance fraud is acceptable, we hope to have taken one step towards understanding (and successfully combatting) insurance fraud.

Appendix: Additional measures and analyses.

Here, we provide all the details we left out of the method and result section in the main text.

Pilot studies: What do people think insurance is?

The size of the Tilburg samples was pre-determined to be 40, because we thought we could survey 40 mall-goers in one day and wanted a student sample of the same size. In the MTURK sample with 104 participants, we also measured attitudes towards fraud with the same scale as we did in Study 1a, as well as which types of insurance each participant had (specifically, whether or not they had health, travel, home-owner, disability, life, auto, any P&C, or other insurance).

Study 1a: Wasted insurance payments and attitude towards insurance fraud.

We did not have specific predictions about the size of the effect, therefore we aimed to have 95% power to find a medium-sized correlation ($r = .3$). According to G*Power we would need 134 people. We recruited 175 U.S. MTURK participants to be able to deal with exclusions.

Materials and procedure. Participants were excluded from the study if they did not successfully pass the IMC. Two hundred and thirty four people started the survey, 197 (84.19%) passed the IMC and 17 did not finish, so we were left with 180 participants with complete data.

We exclude 23 (12.8%) people who take less than 20 seconds on either the fraud or entitlement questions. Note that we also only report the data from the remaining 157 participants in Table 1. The results do not meaningfully change if the analyses are run on the full (N=180) dataset.

Study 1b

We aimed to have 95% power to find a correlation as large as we did in Study 2a ($r = .282$) which means we needed 153 people. We recruited 175 U.S. MTURK participants to be able to deal with exclusions.

Participants. People could only participate if they had an approval rate that was greater than 95% and if they lived in the U.S.

Materials and procedure. Participants were excluded from the study if they did not successfully pass the IMC. One hundred and eighty nine people started the survey, 179

(94.71%) passed the IMC and 1 did not finish, so we were left with 178 participants with complete data.

Upon completion of the measures reported in the main text participants also rated to what extent they liked the six features of insurance (1 = Do not like at all, 7 = Like very much). Unsurprisingly, for each feature, liking is significantly correlated with the extent to which it makes insurance feel like a waste of money, r's = .487 - .696. The correlations are positive because of how the variables are coded but they can be interpreted as follows: the more something feels like a bad investment, the less it is liked.

Results

Due to a programming error, we did not record the time participants spent on each page. We thus do not remove data from the analyses based on the amount of time people took to complete the questions.

Study 2a: Travel insurance premium (€65) and attitude towards fraud

Method

Participants. The session in which this experiment was programmed ran for 2 days, and we aimed to have 150 participants. Due to a programming error, the data from the first day (+/- 100 participants) were unusable. Therefore, we only report the data from the second day, on which 52 people showed up to participate.

Materials and procedure. We did not collect any additional measures.

Results

We did not exclude participants and did not run additional analyses.

Footnote Study: Travel insurance premium ($65) and attitude towards fraud

Method

We aimed to recruit 100 people for this study because we thought 50 people per condition was enough. In retrospect, we determined we had 95% power to find a Cohen's d of 0.73 and 80% power to find a Cohen's d of 0.57.

Participants. One hundred and one participants (M_{age}= 33.41, range 18-67; 53 female) completed this study on MTURK in exchange for $0.10 in June 2012. People could only participate if they had an approval rate that was greater than 95% and if they lived in the U.S.

Materials and procedure. The study was programmed in Qualtrics and included an IMC. One hundred thirty one people started the survey, 105 passed the IMC, and we only analyze the data from 101 participants who completed the survey.

The scenario, manipulation, and DV were the same as in Study 2a except for two differences. Instead of 'You pay 65 Euros per year for the insurance', the text was:

You pay $65 per year for the insurance.

In addition, this study was run in English whereas Study 2a was run in Dutch. Upon completion of the survey, participants answered the following question: 'If you've seen the question on reporting the camera as stolen in other surveys, please tick the box below'. Finally, participants indicated their gender, age, and WorkerID, and were allowed to leave a comment if they wanted to.

Results.

Eight people indicated they had seen the question in other surveys and five people took less than 10 seconds to complete read the scenario and complete the DV. We exclude these people and are left with data for 89 participants.[33]

A *t*-test ($t(87) = 0.70$, $p = .485$, $d = 0.15$) and a Mann-Whitney *U*-test ($U = 1083.50$, $z = 0.82$, $p = .412$, $r = .08$) indicate that participants in the premium condition ($n = 42$; $M = 5.29$, $SD = 1.53$; *Mean Rank* = 47.30) did not think it was more acceptable to report the camera as stolen than participants in the control condition ($n = 47$; $M = 5.06$, $SD = 1.45$; *Mean Rank* = 42.95).

Initially, we took this as a 'failure to replicate' the pattern in Study 2a. However, we found that €65 is a relatively high price for annual travel insurance in the Netherlands while $65 is a very good deal for annual travel insurance in the U.S. More succinctly: $65 is cheap. We thus had (at least) two possible explanations for not replicating the finding in Study 2a. First, presenting the premium does not affect the answers on the DV and the marginal effect in Study 2a was a Type-1 error. For the second explanation, please recall that we think a reminder of how much insurance costs triggers the feeling that money spent on insurance is wasted, and that this feeling leads to greater perceived acceptability of fraud. Presenting a policy that costs $65—a good deal—might not have led people to think the money spent on that insurance was wasted, and that therefore, we found no effect the manipulation. Following this reasoning, we predicted that if the price of insurance were higher, we would find an effect of presenting the price on how acceptable people think fraud is.

To test this second explanation, we based the premium in Study 2b on the current price of travel insurance policies (see main text).

[33] Including the data from these people does not meaningfully change the results.

Study 2b: Travel insurance premium ($150) and attitude towards fraud

Method

In this study we wanted to have 95% power to find an effect at least as large as we did in Study 2a (d = 0.53), for which we would need 188 people. However, effect sizes are generally overestimated so we recruited 300 people, which meant we had 95% power to find a Cohen's d of 0.41, and 80% power to find a Cohen's d of 0.32.

Participants. People could only participate if they had an approval rate that was greater than 95% and if they lived in the U.S. We initially had 306 observations but three people took the survey twice so we excluded their data from the analyses.

Materials and procedure. The study was programmed in Qualtrics and, instead of at the beginning, we included an IMC at the end of the survey. Payment was not affected by performance on the IMC and we only checked whether people who did not pay attention responded differently to the manipulation.

This study was very similar to Study 2a but we added: '$150 dollars for annual travel insurance seems' to which participants could respond by ticking one of three boxes, 1 = A low price, 2 = A normal price, 3 = A high price. Finally, participants indicated their gender, age, and WorkerID, and were allowed to leave a comment if they wanted to.

Results

Seventy nine percent of participants passed the IMC successfully and the performance did not affect how acceptable people perceived fraud to be ($F(1, 271) = 1.64$, $p = .202$, $\eta^2 = .006$), and there was no interaction with the manipulation, ($F(1, 271) = 0.03$, $p = .856$, $\eta^2 < .001$). The effect of presenting the premium on how acceptable participants fraud perceived to be remained significant ($F(1, 271) = 5.18$, $p = .024$, $\eta^2 = .019$).

One hundred forty seven (53.5%) thought $150 was a normal price for annual travel insurance and only 41 (14.9%) and 87 (31.6%) thought it was a low or a high price, respectively. The perceived height of the price was not related to how acceptable people perceived fraud to be ($F(1, 269) = 1.56$, $p = .212$, $\eta^2 = .011$), and there was no interaction with the manipulation, ($F(1, 269) = 0.05$, $p = .948$, $\eta^2 < .001$). The effect of presenting the premium on how acceptable participants fraud perceived to be remained significant ($F(1, 269) = 6.32$, $p = .013$, $\eta^2 = .023$).

Study 3: Health insurance premium ($3500) and attitude towards fraud

Method

In this study we wanted to have 95% power to find an effect that resembled the effect sizes in Study 2a ($d = 0.53$) and Study 2b ($d = 0.36$). We decided to take their average ($d = 0.445$) for the calculations, which meant we would need 266 people. However, we again recruited 300 people to be able to deal with exclusions.

Participants. People could only participate if they had an approval rate that was greater than 95% and if they lived in the U.S.

Materials and procedure. Payment was not affected by performance on the IMC and we only checked whether people who did not pay attention responded differently to the manipulation.

Results

One hundred eighty three participants (77.2%) passed the IMC successfully and the performance did not affect how acceptable people perceived fraud to be ($F(1, 233) = 0.36$, $p = .550$, $\eta^2 = .002$). This time we did find an interaction with the manipulation such that the positive effect of presenting the premium on how people thought about fraud was stronger for those who failed the IMC, ($F(1, 233) = 5.15$, $p = .024$, $\eta^2 = .022$). We also still found the main effect of presenting the premium such that people in the premium condition thought fraud was more acceptable ($F(1, 233) = 11.92$, $p = .001$, $\eta^2 = .049$).

In accordance with our intuition, most people (67.5%) thought $3500 was a high price for health insurance, and only 3% and 29.5% thought it was a low or normal price, respectively. The perceived height of the price was not related to how acceptable people perceived fraud to be ($F(2, 231) = 1.36$, $p = .257$, $\eta^2 = .012$), and there was no interaction with the manipulation, ($F(2, 231) = 0.32$, $p = .969$, $\eta^2 < .001$). The effect of presenting the premium on how acceptable participants fraud perceived was also no longer significant ($F(1, 231) = 1.95$, $p = .164$, $\eta^2 = .008$).

If we include the data from all participants who finished the survey the main effect of perceived height and its interaction with the manipulation are non-significant, but the effect of presenting the premium on perceived acceptability of fraud is significant, $F(1, 295) = 5.53$, $p = .019$, $\eta^2 = .018$).

Chapter 5 - Good students insure themselves against failure

Abstract. Insurance generally provides security in financial contexts but sometimes also in non-financial contexts. We test whether adverse selection, propitious selection, or some other type of selection occurs in a non-financial setting: University students had the opportunity to insure themselves against a failing grade on a final exam by attending tutorials and doing homework assignments. By obtaining a so-called insurance point, students who score a 5—which would normally lead them to fail the exam—will get a 6 and pass the exam. The insurance-point does not affect students with grades other than 5. The data we observe are most consistent with propitious selection and we discuss how these findings relate to the traditional insurance literature on selection problems.

Author Note: We thank Ellen Evers for pointing us to the unusual structure of the course and Andries van de Ark for allowing us to collect data in his course.

This chapter is based on: Van Wolferen, J., Inbar, Y., & Zeelenberg, M. (2014). Good students insure themselves against failure.

When people face uncertainty, they sometimes have the option to buy insurance. By paying a small insurance premium, people eliminate the possibility of suffering a larger loss. Generally, insurance provides financial compensation in case something goes wrong, and people buy health, car, travel and other types of insurance to reduce their financial risk. However, there are also many non-financial decisions that share the same basic structure as insurance, namely taking a small, certain loss to avoid a larger, uncertain one. Vaccination is unpleasant but eliminates the possibility of contracting serious diseases. Condoms reduce pleasure but also reduce the likelihood of pregnancy and sexually transmitted infections. Helmets and seatbelts are cumbersome but reduce the chances of serious injury from an accident. Here, we examine another situation of this type, in which university students could insure themselves against a failing grade on a final exam. This situation provides an excellent opportunity to test whether theories on how people respond to financial insurance also apply to non-financial decisions with the same basic structure.

One of the most important topics in the insurance literature is adverse selection, which entails that people who are most at risk are most likely to buy insurance. Adverse selection is the result of asymmetric information in the insurance market: Consumers (the policyholders) are better able to judge their risk than the insurance company is. Consequently, insurers cannot adjust premiums perfectly to the risk level of the (prospective) policyholders. When this is the case, insurance is a better deal for, and hence more attractive to, someone who is more at risk. Insurers do sometimes attempt to counteract adverse selection—for example, by charging higher auto insurance premiums for young men, who are more at risk of an accident. But most often, such segmentation of the market is not possible, since the risk level of individual consumers is not easily observed. Hence, insurance will be overpriced for a substantial part of the population. This may lead low-risk consumers to not buy or opt out of insurance, which forces insurance companies to raise premiums even further in order to cover the expenses of the high-risk consumers. This self-perpetuating process ultimately leads to market-failures (Akerlof, 1970; Arrow, 1963; Pauly, 1968; Rothschild & Stiglitz, 1976), which is why it has received so much attention in the insurance literature.

Testing for adverse selection typically means testing whether there is a positive association between risk level and the level of coverage at the level of the individual consumer. However, a recent review of the empirical literature on this topic indicates that one should not necessarily expect adverse selection in every insurance context (Cohen & Siegelman, 2010). In fact, the opposite of adverse selection—propitious selection—may occur as well. Propitious selection entails that low-risk consumers are more likely to buy insurance than high-risk consumers (Hemenway, 1990). This results in a situation where those who are most likely to take precautions—i.e., to lower their risk level—are also most likely to buy insurance. In these cases, risk aversion may lead to a preference for precaution to reduce the probability of misfortune, as well as to a preference for insurance to reduce the size of the loss if risk materializes. For example, people who are generally cautious might be more likely to buy insurance against burglary *and* more likely to install better locks (thus reducing the likelihood that they need to file a claim).

Insight into the conditions under which adverse selection does and does not happen is crucial for our understanding of how people respond to insurance. Testing for adverse selection in an unusual context forces us to think about the parallels we can and cannot draw between traditional insurance contexts and the one we describe here. In this article we first describe the context in which we collected the data, then we explain the structure of the data and describe the results of our analyses. In the discussion section we return to the specifics of the examined insurance context and elaborate on how these relate to traditional insurance, and how they may explain the observed pattern of results.

Selection in an educational context: the "insurance point"

We examine adverse selection, propitious selection, or a combination of the two in students' decisions to obtain an "insurance point" for a final exam of a required course in Tilburg University's psychology bachelors program. In the course on "Test theory and test diagnostics"[34] second-year psychology majors attend large lectures with approximately 300 students. In addition to these lectures they can choose to attend small-group tutorials. Critically, when we collected our data, attendance at these tutorials was optional. Students could obtain an insurance point[35] by attending all tutorials and successfully completing all optional homework assignments, and they were free to choose whether or not they wanted to get the insurance point.

In the Netherlands, students fail an exam if they score anything between a 1 and a 5, and pass the exam if they score anything between a 6 and a 10. The insurance point in this course is a way for students to insure themselves against scoring a 5 and failing the exam. If students score a 5 on the final exam but have the insurance point (because they attended all tutorials and did all their homework), the 5 will be changed to a 6 (the lowest passing grade). However, if they score a 4 or lower they fail the exam, even if they have the insurance point. Similarly, if they have the insurance point and score a 6 or higher, they keep their original

[34] info on current contents at http://mystudy.uvt.nl/it10.vakzicht?taal=E&pfac=FSW&vakcode=500154

[35] The course instructor referred to it as the "bonus point" but "insurance point" better describes its purpose.

score and pass the exam. In short, the insurance point only makes a difference for students who score a 5.

To obtain the insurance point students had to attend eight two-hour tutorials. For every tutorial, they had to do several assignments that were related to the topics discussed in the previous meeting as well to the topics to be discussed in the upcoming one. At the beginning of each tutorial, the teacher checked whether each student's homework was sufficient or not. Students were allowed an insufficient score only once out of the eight tutorials, or they would not get the insurance point.

Adverse selection predicts that the students who are at risk of scoring a 5 are most likely to obtain the insurance point. Propitious selection predicts that students who are likely to score high are more likely to obtain the insurance point.[36] For both these selection processes to occur, students should have (private) information to estimate their risk of scoring a 5. We think it is plausible that they do because this course is not the first statistics/methods course that the students take. Students take two statistics course in their first year (Statistics-A and Statistics-B), one in their second-year (Statistics-C) and one simultaneously with the current course (Statistics-D). Third-year students who repeat the course could also have taken or could be taking another course (Statistics-E). We think that students can plausibly infer their risk of scoring a 5 from their past and current performance on the other statistics courses. Indeed, we find that the grade obtained for the current course is reliably predicted by past performance. On average, students score a little more than half a grade lower on the current course than their average final scores for their previously taken statistics courses.

Possible patterns in the data

We test the relationship between previous performance on statistics courses and the likelihood that a student obtains the insurance point. Adverse selection predicts that we find an inverted-u shaped relationship between past performance and the likelihood of obtaining the insurance point. Specifically, students who scored somewhere around 5 and 6 on previously taken courses should be most likely to obtain it. Adverse selection would also predict that the probability of obtaining the insurance point is close to zero for the students at the extremes of the past performance distribution.

[36] Note that while 'adverse' and 'propitious' imply 'bad' and 'good', respectively, we do not think either type of selection is preferable here. We use adverse selection to describe that high-risk students are most likely to get the insurance point; we use propitious selection to describe that low-risk students are most likely to get the insurance point.

Propitious selection predicts a linear or curvilinear relationship between past performance and the likelihood of obtaining the insurance point. Specifically, students who scored high (7 and higher) on previous statistics courses should be most likely to obtain it. Propitious selection would also predict that the probability of obtaining the insurance point is low (close to zero) for the students who are most at risk of scoring a 5 (i.e., those who scored somewhere around 5 and 6 on previously taken courses).

In case adverse selection and propitious selection do not occur, we should find no relationship between past performance on statistics courses and choosing to obtain the insurance point. In the current situation, the decision whether or not to obtain the insurance point is also likely to be motivated by factors other than averting the risk of failing the exam by scoring a five. Students may choose to attend the tutorials as a commitment device for studying, they may be interested in the course's contents, or they may attend because they think the tutorials teach them skills they would not acquire by attending the main lectures. It is thus unlikely that we observe a pattern that resembles *only* adverse selection or *only* propitious selection.

Method

On May 25, 2011, 354 2nd year psychology majors took the exam for the course "Testtheorie and Testdiagnostiek" (Test Theory and Test Diagnostics) at Tilburg University. In one of the lectures during the course, the first author announced to students that they would be asked to participate in a study and explained what data would be collected. The data we present here were collected through a questionnaire that participants filled out after they took their exam. For the students that gave us permission, we later collected data on their prior grades through the university's grade database. We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants. Three hundred and eleven students (out of 354, 87.9%) filled out the survey we distributed after the exam and 276 (78.0%) gave us permission to look up their previous grades. Due to missing values, and because we did not have access to some of the students' grades, we use the data from the 268 people of whom we were able to collect their grades on previous statistics courses.

Data collection. The questionnaire was provided along with the exam and students were told that they could voluntarily participate by filling out the questionnaire after they had taken the exam.

On the questionnaire, participants indicated whether or not they had obtained the insurance point. If they did not obtain the insurance point (yes/no), they indicated whether they had tried to do so (yes/no), and if yes, they were asked to provide an open-ended response explaining why they had failed.[37] Participants then indicated how many hours they had spent on the course, and how many hours they had spent studying for the exam. We also collected a subjective measure of effort by asking "How hard did you feel you had to study for this exam in comparison to other statistics courses?" on a scale that ranged from 1 = less hard, to 5 = equally hard, to 9 = harder. In addition, we asked, "How much effort did you put into studying for this exam in comparison to other statistics courses?" on a scale that ranged from 1 = less, to 5 = equal, and 9 = more.

The questionnaire asked participants to provide their 6 digit student identification number, which would allow us to look up their previous grades on the five statistics courses (MTO-A through E) that they might had taken, as well as their unweighted and weighted (by credits) mean grade of all the courses they had taken thus far. If they agreed, we collected their final grade for each of the five statistics courses, even if they had failed the course. In addition, we recorded how often they had taken the exam for each of those courses.

By filling out the survey, participants also agreed that the instructor of the course would share their performance on the current exam, as well as whether or not they had obtained the insurance point.

Results

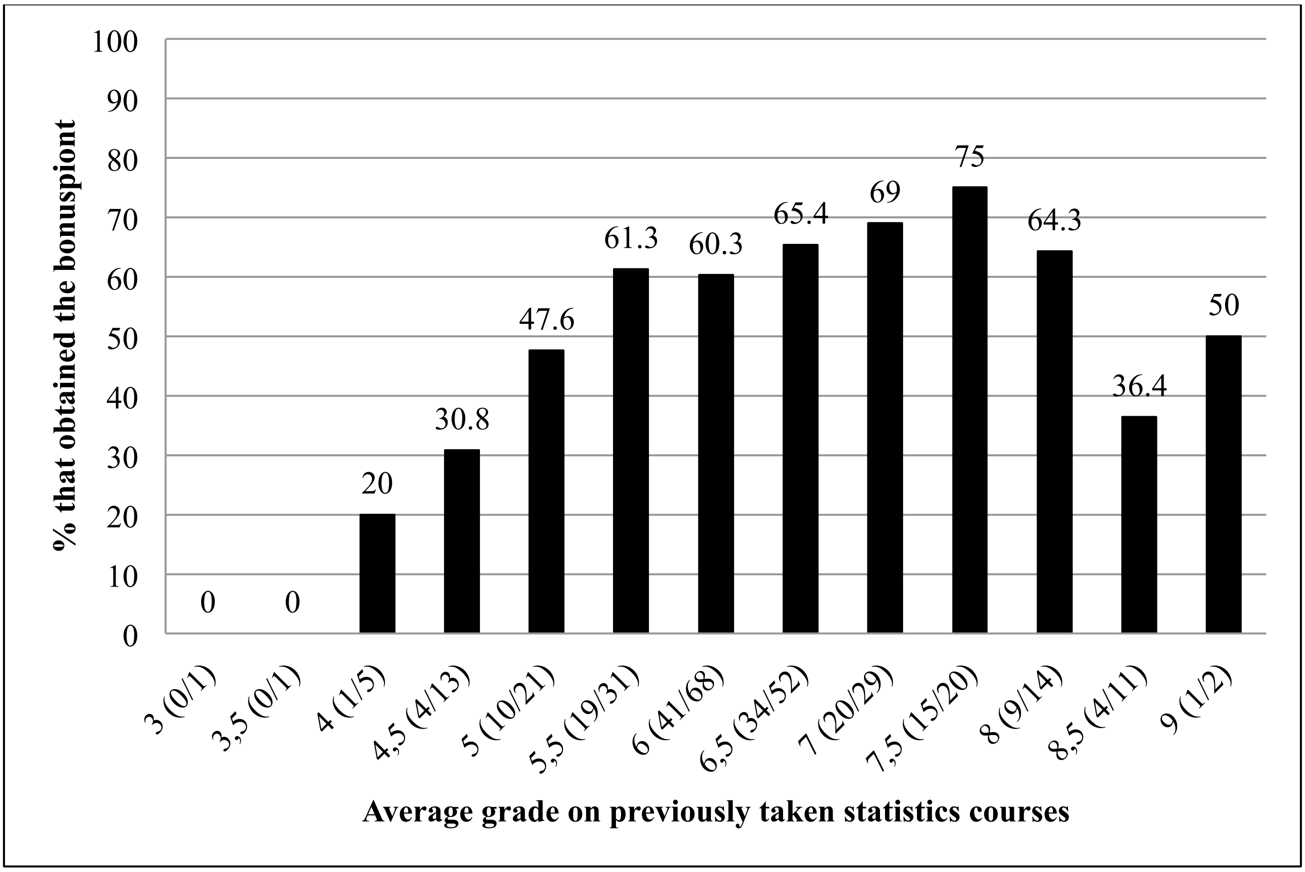
Table 1 provides information about means, percentages, as well as the number of observations and missing values for each variable. To test for the occurrence of adverse or propitious selection effects we need to find which students are at risk of obtaining a five on the final exam. To this end, we test how the average performance on previously taken statistics courses predicts grades on the final exam for this course. We find that on average, people score a half grade lower than they do on average for the other statistics courses (intercept = -0.57, t = -0.98, p = .328, $b_{\text{avgstatsgrade}}$= 0.99, t = 11.22, p < .001). If adverse selection holds, people who score around a 5 and a 6 should thus be most likely to obtain the insurance point. Figure 1 displays the relationship between average score on previous statistics exams and whether or not the student obtained the insurance point.

[37] We coded those answers and assigned them to one of three categories: 1 = self-chosen (e.g., felt like they did not need it), 2 = because of external factors beyond their control (e.g., death in the family). 3 = unclear. However, we had to few observations in each category to use this variable in our analyses.

Table 1

Variable	# valid cases (out of 268)	n (%) 'yes' or Mean (SD)	Observed Min-Max
Insurance point (1 = yes)	267	159 (59.3%)	0-1
Final grade (before insurance point)	267	5.83 (1.85)	1-10
Final grade (after insurance point)	267	5.93 (1.83)	1-10
Statistics A: course taken (1 = yes)	268	257 (95.9%)	0-1
Final grade if taken	257	6.81 (1.01)	4-10
Statistics B: course taken (1 = yes)	268	257 (95.9%)	0-1
Final grade if taken	257	6.75 (1.45)	2-10
Statistics C: course taken (1 = yes)	268	258 (96.3%)	0-1
Final grade if taken	258	5.96 (1.47)	1-9.5
Statistics D: course taken (1 = yes)	268	80 (29.9%)	0-1
Final grade if taken	80	7.02 (1.83)	1-10
Statistics E: course taken (1 = yes)	268	73 (27.2%)	0-1
Final grade if taken	73	4.99 (2.01)	1-10
Avg grade statistics courses	268	6.43 (1.05)	3-9.33
Avg grade all courses taken	268	7.04 (0.46)	6.25-8.70
Avg weighted grade all courses taken	268	7.03 (0.46)	6.25-8.74
Hours spent on course	252	58.54 (46.32)	0-360
Hours of preparation for exam	260	28.23 (25.99)	0-200
Studied hard	264	6.30 (1.52)	1-9
Put in effort	264	5.81 (1.92)	1-9

Figure 1. The percentage of people who obtained the insurance point by how well they previously performed in statistics courses.



Note: The horizontal axis displays the average grade on previously taken statistics courses in 0.5 wide chunks, as well as how many people are in each category and how many of them had obtained the insurance point. For example, 52 people scored between 6.5 and 6.99 on average on previously taken statistics courses, and 34 of them (65.4%) obtained the insurance point.

The pattern we observe is not in line with strong adverse selection, nor with strong propitious selection. Propitious selection would predict a linear relationship between previous performance and probability of obtaining the insurance point. However, if there were some adverse selection, we would predict an inverted u-shape (and thus curvilinear) relationship such that people with higher grades are less likely to obtain the insurance point compared to people with mediocre grades. In our regression models (Table 2) we therefore test for both possibilities. The simplest model, Model 1, only tests for a linear and curvilinear relationship between past performance on statistics courses and the probability of obtaining the insurance point, and finds evidence for both. In Model 2, we include past performance on all courses taken to control for overall ability of the students, as well as self-reported hours spent on the course and on preparation in the exam. We find that the linear and curvilinear relationships still hold.

Inspecting Figure 1, we see that the inverted u-shape we observe is not the kind that is predicted by adverse selection. Students who on average scored between 7.5 and 7.99 on previously taken statistics courses are most likely to have obtained the insurance point (15 out of 20 did so) and that is more consistent with propitious selection. However, propitious selection would predict that the better a student's past performance, the greater his or her likelihood of obtaining the insurance point. The curvilinearity in our data seems to be at least partially driven by the few students at the high end of the past performance distribution. They are *less* likely to have obtained the insurance point. So although propitious selection seems to be the most appropriate description of the selection that occurs in this setting, it does not apply to the highest end of the performance-distribution.

Table 2. Logistic regression models with dependent variable whether or not a student has obtained the insurance point (0 = no, 1 = yes).

Variable	Model 1 *B* (SE) Exp(B)	Model 2 *B* (SE) Exp(B)
Intercept	-13.24 (3.19) 0.00**	-19.38 (6.33) 0.00**
Avg grade statistics courses	3.99 (1.19) 53.92**	4.78 (1.42) 118.78**
(Avg grade statistics courses)^2	-0.28 (0.09) 0.75**	-.362 (0.11) 0.70**
Avg grade all courses taken		0.58 (0.47) 1.79
Hours spent on course		0.02 (0.01) 1.015**
Hours of preparation for exam		-0.02 (0.01) 0.98*
R^2	0.09	0.14
Model test	$\chi^2(2) = 17.30$, $p < .001$	$\chi^2(5) = 26.95$, $p < .001$

Note.* <.05, ** <.01

Discussion

The unusual possibility of obtaining insurance against failing the final exam for a course at our university provided the opportunity to test for adverse selection (Rothschild & Stiglitz, 1976) and propitious selection (Hemenway, 1990) in a non-financial insurance context. Adverse selection predicted that the students who were most likely to need the insurance point would be most likely to have obtained it. Propitious selection predicted that higher-scoring students would be more likely to have obtained the insurance point. The data are most consistent with propitious selection. In this section we discuss which factors may have influenced students' decisions, and attempt to explain why we observe propitious selection rather than adverse selection. In doing so, we follow relevant suggestions by Cohen and Siegelman (2010) and focus on (1) whether or not students had useful private information and whether that information is equally available to high- and low-scoring students, (2) whether or not students use that information, and (3) whether risk and risk-aversion are correlated such that they produce propitious selection.

Which students, if any, have useful (private) information? In order to determine whether they need the insurance point, students need to be able to estimate the probability that they would score a five on the final exam. We found that the average performance on previously taken statistics courses was a strong predictor of the grade students obtained for the current course. We concluded that students could reasonably infer their risk of scoring a five using this proxy, but we did not take into account the uncertainty in each individual's estimate. That is, some students may have scored a four and an eight, while other students may have scored a six twice. The first group of students has more uncertainty regarding their performance than the second and, if they are risk-averse, they should be more willing to get the insurance point than the second group. To test whether this uncertainty predicts uptake of the insurance point, we calculated the standard deviation of the average performance on previously taken statistics courses and added it in the regression model. Unsurprisingly, we find that average performance and the standard deviation around that average performance are negatively correlated, $r = -.36$, $p < .001$. So students who score lower on average have greater uncertainty about their skill (larger *SD* of performance). However, by itself, uncertainty about past performance does not predict the likelihood of obtaining the bonus point (intercept = -0.61, $t = 5.25$, $p = .022$, $b_{\text{SDstats}} = -0.23$, $t = 0.94$, $p = .331$), and it also does not explain any additional variance compared to Model 1 in Table 1, $\chi^2(1) = 0.12$, $p = .726$. It thus seems that students do not use the uncertainty about their performance in their decision to obtain the insurance point.[38]

If only a subset of the population has private information the correlation between risk and uptake of the insurance point may be suppressed if the subset without information is sufficiently large (Cohen & Siegelman, 2010). However, there is no reason to believe that information about past performance is not available to some students. All students know their grades and it is unlikely that they forget how poor/well they did on previous tests.

Do students use the information about their skill? Based on the previous paragraphs, the answer seems to be: no, at least not in the manner that adverse selection predicts. Perhaps this occurs because the students who need the insurance point the most are also the ones that have the most uncertainty regarding their skill. The average standard deviation for people scoring six or lower is 1.21 (*SD* = 0.60) while it is 0.88 (*SD* = 0.47) for people scoring higher than six. This greater uncertainty may lead students to discount the usefulness of their perceived ability and focus on other factors to decide whether or not to get the insurance point.

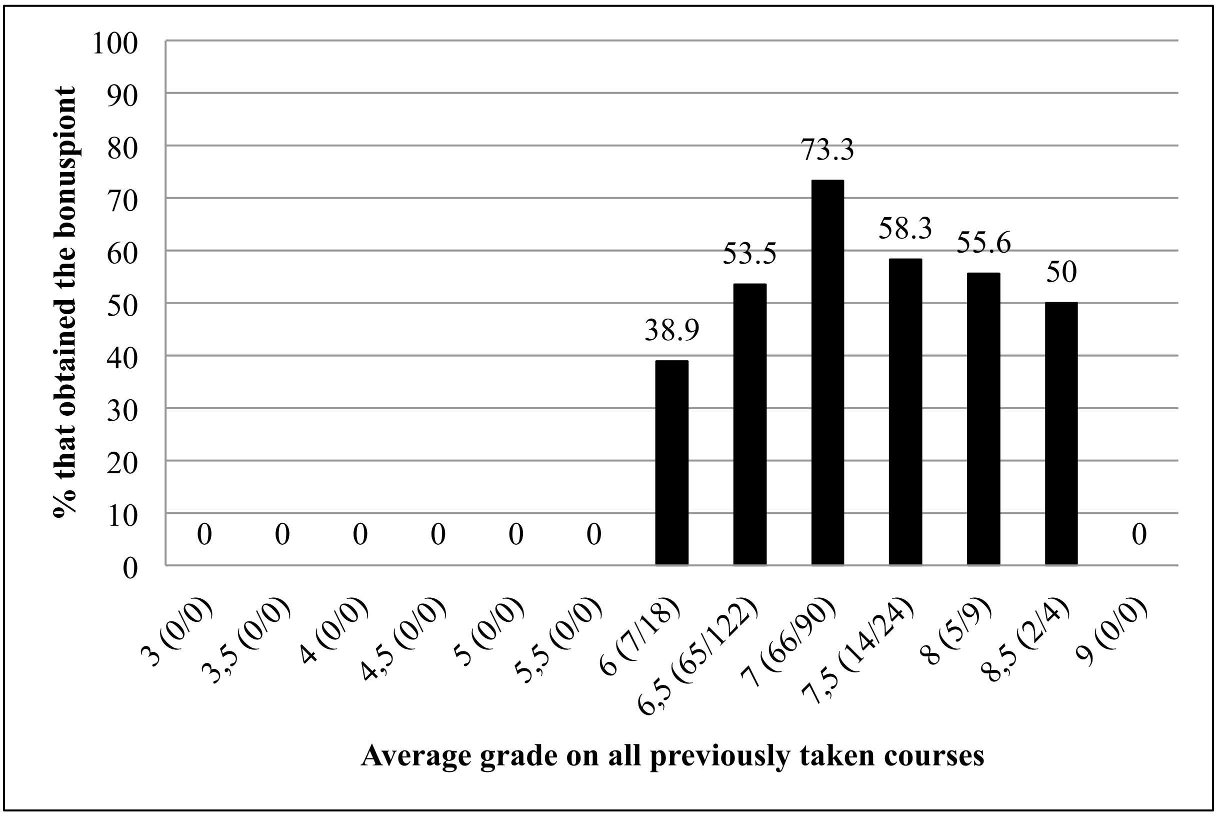
As mentioned in the introduction, there are other factors that students take into account when deciding whether or not to obtain the insurance point. We have no information about their motives and we have no data that we can use to infer their motives. All we can say is that those other factors will obscure possible patterns that are due to adverse, propitious, or other types of selection.

One alternative explanation for the pattern we observe may have come to mind: good students study hard. Studying hard means attending all tutorials and doing all the homework, which would lead students to 'accidentally' obtain the insurance point. We do not think that this is a parsimonious explanation for the findings because this reasoning would predict that the students who scored the highest are most likely to have obtained the insurance point. Figure 1 clearly shows that this is not the case and the significant curvilinear effects in Model 1 and 2 also speak against this reasoning. As a robustness check, we also plotted the probability of obtaining the insurance point depending on previous performance on all courses taken (see Figure 2). The 'good students study hard' explanation would again predict that the highest scoring students are most likely to have obtained the insurance point, but this is not

[38] A different interpretation is that students may be better able to estimate their risk than we did because we estimate uncertainty very crudely based on maximally five grades (mode = 3). We cannot rule out that students in fact do use the uncertainty about their skill in determining whether or not they should obtain the insurance point. However, the fact that we do not find (strong) adverse selection speaks against this interpretation.

what we find. In a regression model predicting whether or not students have the insurance point, we find a significant positive effect of previous performance (intercept = -66.1, t = 9.38, p = .002, $b_{allcourses}$= 17.91, t = 9.15, p = .002), but also a significant curvilinear effect ($b_{allcourses}{}^2$ = -1.20, t = 8.75, p = .003). This suggests that the inverted u-shape in Figure 2 can be interpreted; higher scoring students are indeed less likely to have the insurance point.
We think that not all students use the private information similarly. The lower-scoring students seem to use their information on how well they are likely to do only to some extent while the intermediate students are too pessimistic about their likely performance. The highest-scoring students, however, may use the information about their skill 'rationally' and think they are unlikely to need the insurance point. In fact, we think that the insurance point may be unappealing to these students because if they scored a six they would retake the exam in order to get a higher grade (in general, students are allowed to retake exams at least once). This suggests that there is no one way that describes how *all* students use their private information.

Figure 2. The percentage of people who obtained the insurance point by how well they previously performed in all courses.



Note: The horizontal axis displays the average grade on all previously taken courses in 0.5 wide chunks, as well as how many people are in each category and how many of them had obtained the insurance point. For example, 122 people scored between 6.5 and 6.99 on average on previously taken statistics courses, and 65 of them (53.5%) obtained the insurance point.

Are risk and risk-aversion correlated to produce propitious selection? We do not have a measure for individual differences in risk aversion (beyond whether or not students obtained the insurance point). To explain the pattern we observe in the data we would have to assume that intermediate skill is associated with the greatest aversion to scoring a five. We think it is reasonable to assume that students who usually score around a 7.0 would be more annoyed by failing the exam than students who usually score around a 6.0. A slightly more strenuous assumption (that we hinted at above) is that the students who score even higher are so averse to failing the exam *and* to passing with a six, that they do not consider the insurance point a satisfying way to pass the course.

Another explanation involves an argument about the ability of students to obtain the insurance point. We saw that past performance on statistics courses predicted performance on the current exam. It is likely that the statistics skills are also positively correlated with the students' ability to pass the homework assignments. If we draw the parallel to the insurance literature, we could say that this is a problem of liquidity constraints. The lower scoring students who need the insurance point cannot 'afford' it. They are unable to successfully complete all homework-assignments and therefore do not manage to obtain the insurance point. The relatively more able students are in a similar manner more likely to perform well and obtain the insurance point.

It thus seems that there is no single explanation that parsimoniously describes how students decide whether or not they obtain the insurance point. Low scoring students possibly want the point but are unable to obtain it, intermediate students are willing and able to obtain the point (but we can question whether they should), and high scoring students are able to get it but don't value the insurance point enough to actually do so.

Taken together, we have studied the "insurance" decisions of psychology students in a statistics class and found that the students that are most likely to take up insurance are unlikely to need it. Students that actually could use the insurance are less likely to take it up. As such, these data thus reveal that in this population, for this specific insurance decision, propitious selection is more likely than adverse selection. In agreement with Cohen and Siegelman (2010), we think these findings confirm that adverse selection should only be expected under certain circumstances, and that it is worthwhile to consider contextual factors when testing for selection effects.

Discussion

This dissertation dealt with the psychology of insurance and more specifically with how people's judgments and decisions are affected by insurance. Here, I summarize the findings before I return to the questions posed in the Introduction. In Chapter 1, I provide an overview of the literature on whether, when, and how moral hazard affects people who have health insurance. In Chapter 2, I replicate two magical thinking studies that test how risk perception changes when people know that they are (not) insured or prepared. I find consistent evidence for the 'tempting fate' effect (Risen & Gilovich, 2008) that not being insured makes misfortune seem more likely. However, I find no evidence for the 'insurance effect' (Tykocinski, 2008, 2013) that being insured makes misfortune seem less likely. In Chapter 3 I find experimental evidence for ex ante moral hazard. The controlled tests I report reveal the nature of the effect of insurance on risk-taking. Specifically, I find that people are less willing to take risk when they do not have insurance compared to when they do have insurance, and compared to a situation in which insurance is not salient at all. In addition, I find that labeling a situation as uninsured induces carefulness, even if the incentives that people face are equivalent to a situation that is labeled as insured. In Chapter 4 I find that people are generally unaware of the risk sharing nature of insurance. I propose that this leads policyholders to believe that filing illegitimate claims is a way of compensating for the perceived imbalance between premiums paid and returns obtained. I find that the more people think the money they spend on insurance is wasted, the more they find insurance fraud acceptable. In addition, people find insurance fraud more acceptable when they are reminded of something that makes them feel like they are wasting money on insurance (i.e., paying premiums). In the fifth and final chapter I test which type of selection occurs for a non-financial type of insurance. The pattern of behavior by students who could obtain insurance against failing their exam was not consistent with adverse selection (e.g., Rothschild & Stiglitz, 1976) but rather with propitious selection (Hemenway, 1990). All these chapters build on knowledge from the behavioral sciences and together, they provide new insights into the psychology of insurance. To put these findings into context, I return to the questions I posed in the Introduction and discuss the answers my research gives. I also elaborate on the new questions my findings spur.

When and why do people opt for insurance?

As described in the Introduction, theories of selection effects in insurance markets have developed much faster than the empirical work that tests those theories (see Dionne, 2000). In Chapter 5 I add to the empirical work on selection effects and find that students' decisions to obtain insurance against a failing grade on their exam fit best with propitious selection. Even though there are many differences between traditional financial insurance contexts and the one studied in Chapter 5, the high uptake of the insurance point makes one wonder *why* people choose to get insurance.

An answer to this question could help explain some of the anomalies described in the insurance literature. Sometimes, people buy too much or too little insurance, or they buy the wrong type. For example, there are young and healthy people who do not have health insurance to cover the costs of catastrophic events, and elderly without dependents who buy life insurance rather than an annuity (for more examples and a discussion of other anomalies see Cutler & Zeckhauser, 2004). Similarly, the students in Chapter 5 who usually scored around 8 in statistics courses were very unlikely to need the insurance point. Yet, more than half of those students got the insurance point.

One explanation that (to my best knowledge) has not been empirically tested is that people buy insurance to have peace of mind. This would fit with the research I describe in my dissertation, as well as the literature on insurance decisions. First, even though the insurance effect (i.e., misfortune seems less likely when people are insured; Tykocinski, 2008; 2013, and Chapter 2) probably occurs under limited conditions only, it suggests insurance is associated with feelings of safety. If people anticipate such an effect, regardless of whether they are correct, they might buy insurance to feel safer. Second, the carefulness by people without insurance I observed in Chapter 3 is consistent with the idea that *not* having insurance makes people feel anxious and unwilling to take risk. Having insurance could alleviate such feelings and enable people to experience the benefits of taking risk. Third, the returns (in the form of payments from the insurance company; see Chapter 4) that people focus on when they think about insurance may provide peace of mind. People know that they do not have to worry about the financial consequences of misfortune when they have insurance. In fact, some of the participants in the pilot studies of Chapter 4 alluded to this feature when they described insurance. Fourth, the low uptake of flood-insurance in high-risk areas might be explained by people's unawareness of the risks they face (also see Kunreuther, 1987). They experience no anxiety that needs relief and they consequently do not buy the insurance policy they need. Although speculative, these four claims all fit with the idea that people buy insurance so they can have peace of mind.

If obtaining peace of mind is truly one of the reasons why people buy insurance, insurance could obviously be marketed as relieving anxiety or providing a sense of safety. More importantly, I think insurance companies should then convince their policyholders that the 'return' they are looking for comes in the form of peace of mind. Marketing campaigns that explain that having insurance means not having to worry could lead to less dissatisfied customers and reduce the feeling that people are wasting money on insurance. Based on the findings in Chapter 4, I would predict that such campaigns ultimately help reduce tolerance towards insurance fraud.

In sum, one reason for buying insurance might be that people are looking for peace of mind. Although I did not empirically test this idea, it is worth exploring in future studies in order to get a(n even) better understanding of when and why people buy insurance.

(How) does insurance affect risk perception?

The studies I report in this dissertation do not support the idea that insurance makes misfortune seem less likely. However, it is difficult to draw conclusions from null-results. I think it is clear that insurance does *not always* lead to decreases in perceived risk, but the boundary conditions of the insurance effect have yet to be identified. I think researchers interested in the insurance effect should not take the findings in Chapter 2 as mere failures to replicate, and the findings in Tykocinski (2008) and Tykocinski (2013) as successes. Rather, it is helpful to look at the variation in effect size across samples to see if there are consistent moderators of the effect (IJzerman, Brandt, & Van Wolferen, 2013). Currently, there are not enough estimates of the effect size to test for the existence of moderators, but there is one interesting possibility that warrants further investigation. Perhaps, the counterfactual (i.e., not having insurance) was more salient in studies that find evidence for the insurance effect than in the studies that do not. A counterfactual puts the current situation in perspective and might lead people to feel lucky or relieved that they have insurance. Applying this thought to the available studies, the train commuters in Tykocinski (2008) might have taken more time to think about what it meant to have insurance than the students, train commuters, and mturkers in Van Wolferen et al. (2013a). Consequently, they might have been more likely to compare being with and without insurance. Similarly, having gas masks is rather unusual so the possibility of not having had those masks might have been very salient for the people who picked up the phone in Tykocinski (2013). One suggestion for research on the insurance effect is thus to focus on whether the (mental) comparison between the insured and the uninsured situation is required for the insurance effect to occur.

The risk-perception studies I replicated in Chapter two focus on the perceived likelihood of the outcome (often a loss). Risk = outcome × probability, and the research so far has not tested effects on the perception of the outcome. Obviously, insurance reduces the financial loss associated with a negative outcome, but what happens to the perception of the non-financial part of loss? Following the risk as feelings literature (Loewenstein, Weber, Hsee, & Welch, 2001) and work on the affect heuristic (e.g., Finucane, Alhakami, Slovic, & Johnson, 2000) one would predict that the feelings of safety lead to smaller perceived losses. The sense of safety associated with insurance that was hypothesized to underlie the insurance-effect should also influence the perception of losses. In other words, the literature predicts that there should be an insurance effect for outcomes, such that negative outcomes seem smaller with insurance than without insurance.

(How) does insurance affect risk taking?

In the section on ex ante moral hazard in Chapter 1 it became clear that there remains uncertainty about the size of the effect of insurance on risk-taking. In addition, it is clear that more research is needed to identify the conditions under which ex ante moral hazard does and does not occur. Since the publication of Chapter 1 (Van Wolferen, Inbar, & Zeelenberg, 2013b) at least two new studies have been reported that test whether health insurance is associated with risk-taking (Jerant, Fiscella, Tancredi, & Franks, 2013; Tavares, 2014). While the 2013 paper finds no evidence for ex ante moral hazard, the 2014 paper does. In Chapter 4, I conducted controlled tests of ex ante moral hazard that revealed the nature of the effect. In our studies, ex ante moral hazard is best explained as carefulness by people without insurance. Overall, people with and without insurance seem to differ a little bit in the extent to which they take risk. However, uncertainty remains regarding the boundary conditions of these effects.

The continued work on moral hazard is valuable because of the implication it has for insurance markets. Precise estimates of ex ante moral hazard effects will help policymakers to predict how policy changes will affect behavior of people with and without insurance. Therefore, more work needs to be done. In the final section of this discussion I propose two ways in which I think moral hazard research should proceed.

Why do people find it acceptable to defraud insurance companies?

Chapter 4 highlights a source for the perceived inequity between the output by insurance companies and the input by its policyholders. The focus on how much money (if any) the insurance company gives back to the policyholders in exchange for the premiums it receives leads people to feel like they are wasting money on insurance, and consequently to think fraud is acceptable. I propose that most policyholders think about insurance in a way that facilitates positive attitudes towards fraud, or that they are at least susceptible to the idea that they are wasting money on insurance because ‘they receive nothing in return’. Therefore, my research paints a grimmer picture of who might be susceptible to committing insurance fraud than some of the other work on insurance fraud. Although general theories of cheating suggest that almost everybody is willing to cheat a little bit (Mazar, Amir, & Ariely, 2008), specific research on insurance fraud suggests otherwise. In particular, one theory suggests that insurance fraud is a self-control problem that is mainly committed by risk-takers who like to gamble, and who also commit other crimes such as tax evasion (Ganon & Donegan, 2006). One major difference between that study and mine is that I use hypothetical decisions and measure attitudes, whereas Ganon and Donegan had access to real-world data on who committed insurance fraud and who did not. In research on the effect of deductibles, the findings in studies that employ hypothetical scenarios and measure attitudes towards insurance fraud (Miyazaki, 2008) converge with findings that observe real claiming behavior (Dionne & Gagné, 2001). In these studies higher deductibles are associated with greater tolerance of fraud as well as with a greater likelihood of fraudulent claims. Even though theses studies indicate that insurance fraud research on hypothetical decisions paints a similar picture as research on real decisions, I think my research could benefit from a study that tests whether the way policyholders think about insurance is associated with the likelihood that they file a (real) fraudulent claim.

Future directions

I think the findings related to ex ante moral hazard in Chapter 1, Chapter 2, and Chapter 3 warrant further investigation. As mentioned previously, I think there are two directions in which research on ex ante moral hazard should head and I hope I get the opportunity to be a part of both. First, it is time for a meta-analysis on the effect of insurance on risk taking. This analysis should combine data from all available insurance domains (life, health, car, travel, etcetera) and for each domain it should identify the behaviors that could be affected by ex ante moral hazard. In addition, it should collect published and unpublished work, including experiments run in (and if available, outside) the lab as well as studies that use field-data. Importantly, it should carefully consider whether the collected papers successfully deal with selection-effects. Such an analysis should provide insight in whether there are structural differences between the situations in which ex ante moral hazard is observed, and the situations in which it is not.

Second, the ex ante moral hazard literature would benefit greatly from field experiments, preferably run in collaboration with insurance companies. Ideally, someone would find the time and money to replicate and extend the RAND-experiment (Newhouse, 1993). However, much smaller experiments would advance the ex ante moral hazard literature too. Insurance researchers should take advantage of the policy changes that are occurring throughout the world rather than waiting another 5 years to retrospectively observe the effect of the current changes (also see the discussion in Chapter 1). Insurance companies should allow researchers to offer a randomly chosen (yet representative) subsample of policyholders to keep their old policy or to offer them even more generous coverage. Such designs eliminate confounding selection effects and provide a clean test of ex ante moral hazard. Another advantage of working with insurance companies from the start is that researchers can influence the type and extent of data-collection during the experiment, rather than having to find out what is available after the fact. Specifically, the experiment should collect detailed data on (changes in) smoking behavior, exercise, alcohol intake, and food intake by having participants fill out daily or weekly diaries. This will provide a precise estimate of ex ante moral hazard effects on different preventive and risk taking behaviors.

Together with the meta-analysis, I think experiments will help identify the conditions under which ex ante moral hazard does and does not occur. Ultimately, this research strategy will provide policymakers with the knowledge they need to make decisions in the best way they can.

References

Abbring, J. H., Chiappori, P. A., & Pinquet, J. (2003). Moral hazard and dynamic insurance data. *Journal of the European Economic Association, 1*, 767-820. doi: 10.1162/154247603322493159

Akerlof, G. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*, 488–500.

Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, *53*, 941-973.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M. van Aken, M. A. G., Weber, H., & Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality, 27,* 108-119.

Baker, T. (1996). On the genealogy of moral hazard. *Texas Law Review*, *75*, 237–292.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323–370.

BBC News. (2006). *Sisters lose second coming cover*. Retrieved March 24, 2014, from http://news.bbc.co.uk/2/hi/uk_news/scotland/highlands_and_islands/5105946.stm

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, *76*, 169–217.

Benartzi, S. & Thaler, R. H. (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives*, *21*, 81–104. doi:10.1257/jep.21.3.81

Bernstein, P. L. (1998). *Agains the gods: The remarkable story of risk*. New York: John Wiley and Sons.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*, 3–5. doi:10.1177/1745691610393980

Button, M., Pakes, F., & Blackbourn, D. (2013). *Profile of a Household Insurance Fraudster* (pp. 1–18). Retrieved from http://www.port.ac.uk/departments/academic/icjs/centreforcounterfraudstudies/documents/filetodownload,172919,en.pdf

Carpenter, S. (2012). Psychology’s bold initiative. *Science, 335*, 1558–1561.

Centraal Bureau voor de statistiek (2012). *Zorgrekeningen; uitgaven (in lopende en constante prijzen) en financiering*. Retrieved from http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=71914NED&D1=025,32,37-45&D2=4-13&HDR=G1&STB=T&VW=T

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*, 1145–1177. doi:10.1257/aer.99.4.1145

Chiappori, P., & Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of Political Economy*, *108*, 56–78.

Coalition Against Insurance Fraud. (1997). *Four faces: Why some Americans do - and do not - tolerate insurance fraud*. Retrieved from http://www.insurancefraud.org/downloads/Four%20Faces.pdf

Coalition Against Insurance Fraud (2012). *Go figure: Fraud data*. Retrieved from http://www.insurancefraud.org/healthinsurance.htm

Coalition Against Insurance Fraud. (2013). *How big is $80 billion?* Retrieved March 25, 2013, from http://www.insurancefraud.org/80-billion.htm#.UVB7J1t35fQ

Convenant aanpak verzekeringsfraude (2011). Retrieved from http://www.verzekeraars.nl/UserFiles/File/download/Convenant_Aanpak_Verzekerinsfraude.pdf

Chiappori, P. A., Durand, F., & Geoffard, P. Y. (1998). Moral hazard and the demand for physician services: First lessons from a French natural experiment. *European Economic Review, 42*, 499-511.

Chiappori, P. A. & Salanie, B. (2002). *Testing contract theory: A survey of some recent work*. CESifo working paper 738.

Clarke, M. (1990). The control of insurance fraud. *British Journal of Criminology, 30*, 1-23.

Cockx, B. & Brasseur, C. (2003). The demand for physician services - Evidence from a natural experiment. *Journal of Health Economics, 22*, 881-913. doi: 10.1016/s0167-6296(03)00047-x

Cohen, A. & Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, *77*, 39–84. doi:10.1111/j.l539-6975.2009.01337.x

Courbage, C. & de Coulon, A. (2004). Prevention and private health insurance in the U.K. *Geneva Papers on Risk and Insurance-Issues and Practice, 29*, 719-727.

Cutler, D. M. & Zeckhauser, R. (2004). Extending the theory to meet the practice of insurance. *Brookings-Wharton Papers on Financial Services*, *1*, 1-53. Retrieved from http://muse.jhu.edu/journals/pfs/summary/v2004/2004.1cutler.html

Dave, D. & Kaestner, R. (2009). Health insurance and ex ante moral hazard: Evidence from Medicare. *International Journal of Health Care Finance & Economics, 9*, 367-390. doi: 10.1007/s10754-009-9056-4

Debebe, Z. Y., van Kempen, L., & de Hoop, T. (2012). A perverse 'net' effect? Health insurance and ex-ante moral hazard in Ghana. *Social Science & Medicine*, *75,* 138-147. doi: 10.1016/j.socscimed.2012.02.035

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, *47*, 315–372. doi: 10.1257/jel.47.2.315

Derrig, R. A. (2002). Insurance fraud. *Journal of Risk and Insurance, 69*, 271-287. doi: 10.1111/1539-6975.00026

Dionne, G. (2000). *Handbook of Insurance*. New York: Kluwer Academic Publishers.

Dionne, G., & Gagné, R. (2001). Deductible contracts against fraudulent claims: evidence from automobile insurance. *Review of Economics and Statistics*, *83*, 290–301.

Dionne, G., Giuliano, F., & Picard, P. (2008). Optimal auditing with scoring: Theory and application to insurance fraud. *Management Science*, *55*, 58–70. doi: 10.1287/mnsc.1080.0905

Dionne, G. & Michel, P. (1991). Workers compensation and moral hazard. *Review of Economics and Statistics, 73*, 236-244.

Ehrlich, I. & Becker, G. (1972). Market insurance, self-insurance, and self-protection. *Journal of Political Economy*, *80*, 623–648.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Feldstein, M. (1996). Social security and saving: New time series evidence. *National Tax Journal, 49*, 151-164.

Feynman, R. P., & Leighton, R. (1997). *"Surely You're Joking, Mr. Feynman!" Adventures of a curious character*. New York: Norton & Company.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*, 1–17. doi: 10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fiske, A. P. (1991). *Structures of social life: The four elementary forms of human relations*. New York: Free Press.

Freeman, J. D., Kadiyala, S., Bell, J. F., & Martin, D. P. (2008). The causal effect of health insurance on utilization and outcomes in adults: A systematic review of US studies. *Medical Care, 46*, 1023-1032. doi: 10.1097/MLR.0b013e318185c913

Ganon, M. & Donegan, J. (2006). Self-control and insurance fraud. *Journal of Economic Crime Management*, *4*, 1–24.

Gee, J., Button, M., Brooks, G., & Vincke, P. (2010). *The financial cost of healthcare fraud* (pp. 1–17). Retrieved from http://eprints.port.ac.uk/3987/1/The-Financial-Cost-of-Healthcare-Fraud-Final-(2).pdf

Gouldner, A. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*, 161–178.

Haase, N., Renkewitz, F., & Betsch, C. (2013). The measurement of subjective probability: Evaluating the sensitivity and accuracy of various scales. *Risk Analysis: An Official Publication of the Society for Risk Analysis*, *7*, 1-17. doi: 10.1111/risa.12025

Harper, R. F. (2010). *The code of hammurabi king of babylon. About 2250 B.C. autographed text transliteration* (Reprint.). Clark, NJ: The Lawbook Exchange, Ltd.

Heiman, G. W. (2002). *Research methods in psychology* (3rd ed.). Boston, MA: Houghton Mifflin Company.

Hemenway, D. (1990). Propitious selection. *Quarterly Journal of Economics*, *105*, 1063–1069.

Hershey, J. & Schoemaker, P. (1980). Risk taking and problem context in the domain of losses: An expected utility analysis. *Journal of Risk and Insurance*, *47*, 111–132.

IJzerman, H., Brandt, M. J., & Van Wolferen, J. (2013). Rejoice! In replication. *European Journal of Personality, 27,* 128-129.

Insurance Information Institute. (2010). *Insurance Handbook - A guide to insurance: what it does and how it works* (pp. 1–205). New York, NY: Insurance Information Insitute. Retrieved from http://www2.iii.org/assets/docs/pdf/Insurance_Handbook_20103.pdf

Insurance Journal. (2012). *U.S. insurers say fraud costs on the rise: Survey*. Retrieved March 26, 2013, from http://www.insurancejournal.com/news/national/2012/10/09/265939.htm

Insurance Services Office. (2000). *Fighting insurance fraud - Survey of insurer anti-fraud efforts*. Retrieved March 25, 2013, from http://www.iso.com/Research-and-Analyses/Studies-and-Whitepapers/Fighting-Insurance-Fraud-Survey-of-Insurer-Anti-Fraud-Efforts.html#

Israel Ministry of Foreign Affairs (2002). *The health care system in Israel - a historical perspective*. Retrieved from http://www.mfa.gov.il/MFA/History/Modern\%20History/Israel\%20at\%2050/The\%20Health\%20Care\%20System\%20in\%20Israel-\%20An\%20Historical\%20Pe

Jenni, K., & Loewenstein, G. (1997). Explaining the identifiable victim effect. *Journal of Risk and Uncertainty*, *257*, 235–257.

Jerant, A., Fiscella, K., Tancredi, D. J., & Franks, P. (2013). Health insurance is associated with preventive care but not personal health behaviors. *Journal of the American Board of Family Medicine, 26*, 759–67. doi:10.3122/jabfm.2013.06.130054

Johnson, E., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, *7*, 35–51.

Juurikkala, O. (2008). Retire early, save little: The ironic disincentives of public pension systems. In Booth, P., Juurikkala, O. & Silver, N. (eds.) *Pension Provision: Government failure around the world* (pp. 75-94). London: The institute of economic affairs.

Keating, D. (1991). Pension insurance, bankruptcy and moral hazard. *Wisconsin Law Review*, 65-108.

Keeler, E. B. (1988). *The demand for episodes of medical treatment in the health insurance experiment*. Santa Monica: Rand Corporation.

Kenkel, D. S. (2000). Prevention. In A. J. Culyer & J. P. Newhouse (eds.), *Handbook of Health Economics* (vol. 1B, pp. 1676-1720). Amsterdam, The Netherlands: Elsevier.

Kirchler, E. (2009). *The economic psychology of tax behavior*. Cambridge: Cambridge University Press.

Klick, J. & Stratmann, T. (2007). Diabetes treatments and moral hazard. *Journal of Law & Economics, 50*, 519-538.

Koç, Ç. (2010). Disease-specific moral hazard and optimal health insurance design for physician services. *Journal of Risk and Insurance*. doi: 10.1111/j.1539-6975.2010.01384.x

Kunreuther, H. (1987). *Disaster insurance protection: public policy lessons*. New York: John Wiley and Sons.

Kunreuther, H., Pauly, M. V., & McMorrow, S. (2013). *Insurance and behavioral economics : Improving decisions in the most misunderstood industry*. Cambridge: Cambridge University Press.

Kusev, P., Van Schaik, P., Ayton, P., Dent, J., & Chater, N. (2009). Exaggerated risk: prospect theory and probability weighting in risky choice. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*, 1487–1505. doi: 10.1037/a0017039

Lloyd's. (2014). *History*. Retrieved March 31, 2014, from http://www.lloyds.com/lloyds/about-us/history

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286. doi: 10.1037//0033-2909.127.2.267

Loubergé, H. (2000). The development of risk and insurance economics: The past 25 years. In G. Dionne (Ed.), *Handbook of Insurance* (pp. 3–33). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Major, J. A. & Riedinger, D. R. (2002). EFD: A hybrid knowledge/statistical-based system for the detection of fraud. *Journal of Risk and Insurance, 69*, 309-324. doi: 10.1111/1539-6975.00025

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*, 633–644. doi: 10.1509/jmkr.45.6.633

Michel-Kerjan, E., Lemoyne de Forges, S., & Kunreuther, H. (2012). Policy tenure under the U.S. national flood insurance program (NFIP). *Risk Analysis: An Official Publication of the Society for Risk Analysis*, *32*, 644–58. doi: 10.1111/j.1539-6924.2011.01671.x

Miyazaki, A. D. (2008). Perceived ethicality of insurance claim fraud: Do higher deductibles lead to lower ethical standards? *Journal of Business Ethics*, *87*, 589–598. doi: 10.1007/s10551-008-9960-4

Newhouse, J. P. (1996). *Free for all? Lessons from the RAND Health Insurance Experiment*. Cambridge, MA: Harvard University Press.

Nussbaum, D. (2012). The role of conceptual replication. *The Psychologist, 25*, 350.

Nyman, J. A. (2004). Is 'moral hazard' inefficient? The policy implications of a new theory. *Health Affairs, 23*, 194-199. doi: 10.1377/hlthaff.23.5.194

Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, *7*, 657–660. doi: 10.1177/1745691612462588

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*, 867–872.

Organisation for Economic Cooperation and Development. (2014). *OECD StatsExtracts > Insurance indicators >Total gross premiums*. Retrieved from http://stats.oecd.org/Index.aspx

Pagan, J. A., Puig, A. & Soldo, B. J. (2007). Health insurance coverage and the use of preventive services by Mexican adults. *Health Economics, 16*, 1359-1369. doi: 10.1002/Hec.1226

Pauly, M.V. (1968). The economics of moral hazard: Comment. *American Economic Review, 58*, 531-537.

Pauly, M. V. & Held, P. J. (1990). Benign moral hazard and the cost-effectiveness analysis of insurance-coverage. *Journal of Health Economics, 9*, 447-461.

Picard, P. (2000). Economic analysis of insurance fraud. In G. Dionne (Ed.), *Handbook of Insurance* (1st ed., pp. 315–362). Boston, MA: Kluwer Academic Publishers.

Richaudeau, D. (1999). Automobile insurance contracts and risk of accident: An empirical test using French individual data. *Geneva Papers on Risk and Insurance Theory, 24*, 97-114.

Rijksoverheid (2012). *Wat is het eigen risico in de zorgverzekering*? Retrieved from http://www.rijksoverheid.nl/onderwerpen/zorgverzekering/vraag-en-antwoord/wat-is-het-eigen-risico-in-de-zorgverzekering.html

Risen, J. L. & Gilovich, T. (2007). Another look at why people are reluctant to exchange lottery tickets. *Journal of Personality and Social Psychology, 93*, 12–22. doi: 10.1037/0022-3514.93.1.12

Risen, J. L. & Gilovich, T. (2008). Why people are reluctant to tempt fate. *Journal of Personality and Social Psychology*, *95*, 293–307. doi: 10.1037/0022-3514.95.2.293

Rothschild, M. & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics, 90*, 629-649.

Rowell, D. & Connelly, L.B. (2012). A history of the term “Moral Hazard”. *Journal of Risk and Insurance*, 1-25. doi: 10.1111/j.1539-6975.2011.01448.x

Savedoff, W.D. (2004). Kenneth Arrow and the birth of health economics. *Bulletin of the World Health Organization, 82*, 139-140.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauerschatka, H., & Simons, A. (1991). Ease of retrieval as information - Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*, 195-202.

Schoemaker, P. & Kunreuther, H. (1979). An experimental study of insurance decisions. *Journal of Risk and Insurance*, *46*, 603–618.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11–36.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359–1366.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012) A 21 Word Solution. *Dialogue, 26,* 4–6.

Skriabikova, O., Pavlova, M., & Groot, W. (2010). Empirical models of demand for out-patient physician services and their relevance to the assessment of patient payment policies: A critical review of the literature. *International Journal of Environmental Research and Public Health, 7*, 2708-2725. doi: 10.3390/ijerph7062708

Slovic, P. (1987). Perception of risk. *Science*, *236*, 280-285.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Preference for insuring against probable small losses: Insurance implications. *Journal of Risk and Insurance*, *44*, 237–258.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1988). Response mode, framing, and information-processing effects in risk assessment. In Bell, D. E., Raiffa, H., & Tversky, A. (Eds.). *Decision making: Descriptive, normative, and prescriptive interactions* (pp. 152-166). Cambridge: Cambridge University Press.

Slovic, P. & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15*, 322–325. doi: 10.1111/j.1467-8721.2006.00461.x

Stanciole, A.E. (2007). *Health insurance and life style choices: Identifying the ex ante moral hazard*. IRISS Working Paper Series. Retrieved from http://econpapers.repec.org/RePEc:irs:iriswp:2007-10

Swiss Re Ltd. (2013). *Sigma: World insurance in 2012*. Retrieved from http://www.tsb.org.tr/images/Documents/sigma3_2013_en.pdf

Sydnor, J. (2010). (Over) insuring modest risks. *American Economic Journal: Applied Economics*, *2*, 177–199. doi: 10.1257/app.2.4.177

Tavares, A. I. (2014). Health insurance and lifestyles. *Applied Economics*, *46*, 1910–1923. doi: 10.1080/00036846.2014.889803

Tennyson, S. (1997). Economic institutions and individual ethics: A study of consumer attitudes toward insurance fraud. *Journal of Economic Behavior & Organization, 32*, 247-265. doi: 10.1016/s0167-2681(96)00904-3

Tennyson, S. (2002). Insurance experience and consumers' attitudes toward insurance fraud. *Journal of Insurance Regulation, 21*, 35-55.

Tennyson, S. & Salsas-Forn, P. (2002). Claims auditing in automobile insurance: Fraud detection and deterrence objectives. *Journal of Risk & Insurance*, *69*, 289–308. doi: 10.1111/1539-6975.00024

Thaler, R. H. & Benartzi, S. (2004). Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*, s164-s187. doi: http://dx.doi.org/10.1086/380085

Trenerry, C. F. (1926). *The origin and early history of insurance including the contract of bottomry*. London: P.S. King & Son.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Tykocinski, O. E. (2008). Insurance, risk, and magical thinking. *Personality & Social Psychology Bulletin*, *34*, 1346–1356. doi: 10.1177/0146167208320556

Tykocinski, O. E. (2013). The insurance effect: How the possession of gas masks reduces the likelihood of a missile attack. *Judgment and Decision Making*, *8*, 174–178.

Van de Voorde, C., Van Doorslaer, E., & Schokkaert, E. (2001). Effects of cost sharing on physician utilization under favourable conditions for supplier-induced demand. *Health Economics, 10*, 457-471. doi: 10.1002/hec.631

Van Vliet, R. C. J. A. (2001). Effects of price and deductibles on medical care demand, estimated from survey data. *Applied Economics, 33*, 1515-1524. doi: 10.1080 /00036840010013 62

van Vliet, R. C. J. A. (2004). Deductibles and health care expenditures: Empirical estimates of price sensitivity based on administrative data. *International Journal of Health Care Finance and Economics, 4*, 283-305. doi: 10.1023/B:IHFE.0000043759.93644.e0

Van Wolferen, J., Inbar, Y., & Zeelenberg, M. (2013a). Magical thinking in predictions of negative events : Evidence for tempting fate but not for a protection effect. *Judgment and Decision Making*, *8*, 45–54.

Van Wolferen, J., Inbar, Y., & Zeelenberg, M. (2013b). Moral hazard in the insurance industry. *Netspar Panel Paper #33*, 1–76.

Verbond van Verzekeraars (2011). *Verzekerd van Cijfers 2011*. Retrieved from http://www.verzekeraars.nl/UserFiles/File/cijfers/VvC2011_totaal.pdf

Verschuur, W. L. G. (1992). *Determinanten van verzekeringsfraude*. Leiden: Werkgroep Veiligheid.

Viaene, S. & Dedene, G. (2004). Insurance fraud: Issues and challenges. *Geneva papers on Risk and Insurance, 29*, 313-333. doi: 10.1111/j.1468-0440.2004.00290.x

Wagenaar, W., Keren, G., & Lichtenstein, S. (1988). Islanders and hostages: deep and surface structures of decision problems. *Acta Psychologica*, *67*, 175–189.

Walster, E. H., Hatfield, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research*. Boston: Allyn and Bacon.

Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15*, 263–290.

Wenzel, M. (2002). Tax compliance and the psychology of justice: Mapping the field. In Braithwaite, V. (Ed.), *Taxing Democracy: Understanding tax avoidance and evasion* (pp. 41–70). Aldershot: Ashgate Publishing Ltd.

Zeelenberg, M. & Breugelmans, S. M. (2008). The role of interpersonal harm in distinguishing regret from guilt. *Emotion*, *8*, 589–596. doi: 10.1037/a0012894

Zweifel, P. & Manning, W.G. (2000). Moral hazard and consumer incentives in health care. In Anthony, J. C. & Joseph, P. N. (eds.), *Handbook of Health Economics* (vol. 1, pp. 409-459). Amserdam: Elsevier.

Dankwoord

Hoewel je in een dankwoord meestal iedereen die er (niet) in staat te kort doet wil ik toch heel graag iedereen bedanken die me de afgelopen 4 jaar geholpen, geïnspireerd, of gemotiveerd heeft.

Marcel en Yoel, bazen, vrienden, jullie zijn fantastisch. Marcel, jij staat altijd voor me klaar met slimme ideeën, tips, en advies. Over onderzoek, maar ook over alles dat niets met werk te maken heeft. Daar heb ik de afgelopen 4 jaar ontzettend van geprofiteerd en ik hoop nog heel lang bij jou aan te kloppen voor een praatje, overleg, of een biertje. Yoel, my English wouldn't have been so idiomatic if it hadn't been for you. You taught me everything by being a super smart, eloquent, and laid-back supervisor. Let's not let the distance keep us from seeing each other.

Irene, Maartje, en Yvette, mijn lieve paranimfen. Irene, met jou is het altijd feest en jij bent er op alle mogelijke manieren voor me. Ik ben gek op je en kan niet wachten om met jou nog meer avonturen te beleven. Wij gaan samen een super mooie en spannende toekomst tegemoet. Maartje, jou ken ik het langst en in alle fases van mijn proefschrift was je een super fijne, gezellige, en leuke kamergenoot. We hebben inmiddels aan anderhalf woord genoeg en dat zal ik missen nu ik je niet meer elke dag zie. Yvette, jij bent de grote zus die ik nooit had. Je bent mijn beste vriendin en zonder jou had ik dit project niet eens overwogen. Dank je voor alles.

Joep, Erik, Rutger, en Jurriaan, bedankt voor jullie vriendschap, luisterende oren en fantastische grappen.

Pap, mam, bruure, bedankt dat jullie er altijd zijn en voor alles dat jullie voor me gedaan hebben.

Alle andere collega's, vrienden, en familie, bedankt voor de inspiratie en afleiding op het werk, thuis, en daarbuiten. Zonder jullie was het maar saai geweest!